UNIVERSITY OF CALIFORNIA

Santa Barbara

Investigating the Mechanisms of *in vivo* Cellular Reprogramming and Transorganogenesis in *Caenorhabditis elegans*

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular, Cellular, and Developmental Biology

by

Erik Anthony Spickard

Committee in charge:

Professor Joel Rothman, Chair

Professor Dennis Clegg

Professor Kathy Foltz

Professor William Smith

December 2019

The dissertation of Erik Anthony Spickard is approved.

_____

Kathy Foltz

_____

Dennis Clegg

_____

William Smith

_____

Joel Rothman, Committee Chair

December 2019

Investigating the Mechanisms of *in vivo* Cellular Reprogramming and Transorganogenesis in *C. elegans*

ACKNOWLEDGEMENTS

I would like to thank Jonathan Fischer and Alec Barrios for their dedication to the RNAi screening effort. Thank you to my advisor, Joel Rothman, for the years of opportunities and mentorship, and to my committee members, Kathy Foltz, Bill Smith, and Dennis Clegg, for their advice and wisdom. I am grateful to all members of the Rothman lab for the camaraderie, feedback, and friendship.

I would like to thank my friends and family for their support and genuine interest in my research. I am thankful for my parents, Lynne and Steve, for their unconditional love and encouragement, and for instilling in me an appreciation for the beauty and wonder of the natural world that has made me into the biologist I am.

Thanks to Jennifer Smith, manager of the Biological Nanostructures Laboratory at UCSB, for RNA sequencing assistance. Some worm strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). This work was supported by grants RO1HD081266 and RO1HD082347 from the National Institutes of Health.

DEDICATION

This work would not have been possible without the continuous support of my wife.

I dedicate this dissertation to Grace, for her unwavering belief in my ability to complete it, and her patience throughout this entire journey. I could not have wished for a more loving and encouraging partner. And to my son, Archer, for the perspective and motivation to take the next step.

VITA OF ERIK ANTHONY SPICKARD

December 2019

EDUCATION

Ph.D. in Molecular, Cellular, and Developmental Biology, December 2019
University of California, Santa Barbara, Santa Barbara, CA

Bachelor of Science in Biotechnology, with Honors, December 2009
University of California, Davis, Davis, CA

Education Abroad Program, Fall 2007
University of British Columbia, Vancouver, Canada

PROFESSIONAL EMPLOYMENT

University of California, Santa Barbara
Teaching Assistant
MCDB 109L: Biochemistry Lab, Spring 2019
MCDB 101A: Genetics (Head TA), Fall 2018
MCDB 112L: Developmental Biology Lab, Winter 2015, 2016, 2017, 2018, 2019
MCDB 1AL: Introductory Biology Lab, Fall 2012, 2014
MCDB 1BL: Introductory Biology Lab, Summer 2014
EEMB 3L: Introductory Biology Lab, Spring 2014
MCDB 126B: Molecular Pharmacology, Winter 2014

Gen-Probe Inc., San Diego, CA
Research Associate I, Dec 2010-Aug 2012

Molecular Response, LLC, San Diego, CA
Purchaser/Accounts Payable Assistant, Sept-Nov 2010
Pharmaceutical Manufacturing Technician, Temporary, Sept-Oct 2010

California Animal Health and Food Safety Laboratory, Davis, CA
Lab Assistant III, April-June 2010

UCDSOM Department of Dermatology, Dr. Liu Lab, Sacramento, CA
Immunology Research Intern, February-April 2010

UCD Department of Animal Sci., Animal Genomics and Biotechnology Lab, Davis, CA
Student Research Intern, March-June 2009

Lawrence Berkeley National Laboratory, Berkeley, CA
Science Undergraduate Laboratory Internship (SULI) Program, June-August 2008

PUBLICATIONS

Spickard, E.A., and Rothman, J.H. (in preparation). An intracellular pathogen-like innate immune response is activated during transorganogenesis in *Caenorhabditis elegans*.

Spickard, E.A., Joshi, P.M., Rothman, J.H. (2018). The multipotency-to-commitment transition in *Caenorhabditis elegans*—implications for reprogramming from cells to organs. FEBS Letters 592, 838-851. https://doi.org/10.1002/1873-3468.12977

Riddle, M., Spickard, E., Jevince, A., Nguyen, K., Hall, D., Joshi, P. and Rothman, J. (2016). Transorganogenesis and transdifferentiation in *C. elegans* are dependent on differentiated cell identity. Developmental Biology 420, 136–147.

ABSTRACTS AND PRESENTATIONS

Temporal transcriptional profiling of transdifferentiation and transorganogenesis in *C. elegans*. Poster, ASCB | EMBO 2018 Meeting, San Diego, CA, December 8-12, 2018.

Transcriptome and functional analysis of ELT-7-directed reprogramming and transorganogenesis of the *C. elegans* gonad and pharynx. Poster, 21st International *C. elegans* Conference, University of California, Los Angeles, June 2017.

In Vivo cellular reprogramming and transorganogenesis in *C. elegans*. Poster, MCDB Graduate Student Symposium, 2017.

Gonads to Guts: Reprogramming an Organ in the Nematode *C. elegans*. UC Santa Barbara 3rd Annual Grad Slam, April 17, 2015. Finalist.

Spickard, E., and L. Gundel. Testing of Low-Cost Samplers for Collection of Particulate Matter from Second Hand Smoke. U.S. Department of Energy Journal of Undergraduate Research. Volume IX, 2009.

AWARDS, HONORS, AND FELLOWSHIPS

Doctoral Student Travel Grant, UCSB Academic Senate, 2018
Doreen J. Putrah Cancer Research Foundation Conference Fellowship, UCSB, 2018
Jane Altman Fellow, UCSB, 2016
CNSI Next Generation Sequencing Award, California NanoSystems Institute, 2015
Molecular, Cellular, and Developmental Biology Departmental Fellowship, UCSB, 2013
Charles Storke Award, UCSB, 2013
Rathmann Fellowship, UCSB, 2012

COMMITTEES

Graduate Union of Molecular Biology Investigators (GUMBI), Webmaster, https://gumbi.lifesci.ucsb.edu, 2014-2018

ABSTRACT

Investigating the Mechanisms of *in vivo* Cellular Reprogramming and Transorganogenesis in *C. elegans*

by

Erik Anthony Spickard

Cellular reprogramming has the potential to treat numerous degenerative diseases and repair injured organs, but safe and effective therapeutic applications require deeper understanding of the molecular and cellular consequences of forced cell fate changes. The model organism *Caenorhabditis elegans* can be used to study cellular reprogramming *in vivo* and features diverse reprogramming-related events, including: the multipotency-to-commitment transition during embryogenesis, naturally occurring transdifferentiation, and forced reprogramming of the germline and differentiated somatic cells. A remarkable finding from a previous study showed that ectopic expression of the transcription factor ELT-7 alone is capable of triggering transdifferentiation of the pharynx into intestine-like tissue, and "transorganogenesis" of the entire uterus into a well-formed intestine-like organ. However, the molecular mechanisms involved in this process are largely unknown.

In this work, I primarily utilize reverse genetic screening and time course mRNA-sequencing methods to probe the ability of ELT-7 to initiate reprogramming and to identify biological processes involved in somatic gonad-to-intestine transorganogenesis. I find that

reprogramming to intestine mirrors embryonic endoderm development, engaging most, if not all, of the endoderm gene regulatory network (GRN). Three chromatin-associated factors are identified that appear to be playing important roles in ELT-7-directed transorganogenesis, suggesting that establishment of proper chromatin architecture may define a permissive cellular context for reprogramming into intestine. Bioinformatics analysis of the ELT-7-directed reprogramming temporal transcriptional profile shows that widespread gene expression changes occur even in some non-reprogrammed tissues, indicating significant cell fate elasticity and suggesting cell type-specific modes for cell fate maintenance. Finally, and somewhat surprisingly, gene expression changes following ectopic ELT-7 expression reveal the activation of an intracellular pathogen response (IPR) that correlates both spatially and temporally with cellular reprogramming and may be promoting somatic gonad-to-intestine transorganogenesis.

This work presents the gene expression dynamics of *in vivo* cellular reprogramming and characterizes the distinct biological processes involved. These findings suggest that the cell type-specific reprogramming ability of ELT-7 may rely on the robust endoderm GRN and require a cellular context that is capable of activating the IPR. These data will enable further targeted research into the mechanisms of transorganogenesis and may lead to the discovery of similar phenomena in *C. elegans* and other organisms.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter One

Cellular Reprogramming of single cells and whole organs in *C. elegans**

*This work has been published as:
Spickard, E., Joshi, P., Rothman, J. (2018). The multipotency-to-commitment transition in Caenorhabditis elegans—implications for reprogramming from cells to organs. FEBS Letters. https://dx.doi.org/10.1002/1873-3468.12977

Introduction

Development begins with a single totipotent cell, the fertilized zygote, poised with the potential to become any type of cell within the organism. Through successive cell divisions, embryonic cells undergo binary or multipotential fate decisions, with progressively more restricted levels of specification. During this process, particular cellular contexts are established that restrict cell fate potential, culminating in cells with fully committed differentiated identities. This progressive restriction of cell fate is driven by the stepwise mobilization of transcription factors that comprise key regulatory nodes in gene regulatory networks.

The transformation of a zygote into a multicellular organism, containing a broad range of specialized cell types, is marked by major developmental transitions that have been illuminated in a variety of animal embryos. These include the oocyte-to-embryo transition, during which the developmental program for oogenesis is replaced by the program to initiate development of a multicellular embryo (Robertson and Lin, 2015). Later, the mid-blastula transition (MBT) (Newport and Kirschner, 1982) is activated, during which (particularly notably in vertebrates) the nuclear DNA:cytoplasm ratio triggers widespread transcriptional activation of a previously largely quiescent nucleus. This event leads to the process that occurs more generally in all embryos: the maternal-to-zygotic transition (MZT) (Baroux et al., 2008; Tadros and Lipshitz, 2009), characterized by activation of the embryonic genome, concomitant with elimination of a substantial fraction of the maternal factors that had been laid down in the oocyte during germline development (Baugh et al., 2003; Langley et al., 2014; Robertson and Lin, 2015; Schier, 2007). These latter processes also correlate with loss of developmental plasticity, whereby the initial ability of cells to adopt a wide range of fates

becomes progressively more restricted. Subsequent to these early events, cells undergo dramatic rearrangements in a further major transition known as gastrulation, which establishes the embryonic germ layers (Nance et al., 2005).

As we highlight in studies of *C. elegans*, by probing cellular plasticity in animal embryos it is also possible to discern another major embryonic transition, the multipotency-to-commitment transition (MCT) (Joshi et al., 2010; Riddle et al., 2013; Riddle et al., 2016). Understanding the MCT is pivotal to our knowledge of cell fate commitment, developmental plasticity, and the ability to reprogram cells. Studies of these critical processes benefit greatly from investigations in intact embryos, which provide an *in vivo* cellular context that is generally absent in most reprogramming studies (Srivastava and DeWitt, 2016; Xu et al., 2015). With the ability to track multiple cell and tissue types at the single cell level, and powerful molecular genetic tools that make it possible to assess both minor and more dramatic changes in cell fate, *C. elegans* is particularly well-suited for such investigations.

A broad spectrum of cellular reprogramming events have been observed in *C. elegans*, spanning all stages of the life cycle, both before and after the MCT, and occurring in many distinct cellular contexts (Fig. 1) (Becker and Jarriault, 2016; Hajduskova et al., 2012; Joshi et al., 2010; Tursun, 2012; Zuryn et al., 2012). This review addresses some of the major findings regarding developmental reprogramming in *C. elegans*, grouped by their cellular context, including forced reprogramming of early blastomeres, reprogramming of the germ line, transdifferentiation and transorganogenesis of somatic cells, and naturally occurring transdifferentiation. While there are no universal principles to these events, we discuss common themes, highlight the aspects that are thus far unique to *C. elegans*, and forecast the

future translational impact of reprogramming studies that might emerge from studies in this model system.

The Multipotency-to-Commitment Transition (MCT)

The ability to differentiate cells corresponding to the three classical germ layer types and the germline is differentially apportioned to a series of six "founder cells," which are established during the first several cell divisions in early *C. elegans* embryos. At the time of their birth, the founder cells appear to be specified to follow particular development pathways, as isolated founder cells that are allowed to develop in culture generate largely the same tissue types that they engender in intact embryos (Fig. 1) (Hashimshony et al., 2015; Levin et al., 2012; Levin et al., 2016; Priess and Thomson, 1987). However, although founder cells and their descendants are specified, they are not committed to those fates and can undergo transdetermination. In fact, many studies have shown that progenitor cells produced well after the birth of the founder cells are multipotent and remain competent to follow different major developmental trajectories characteristic of all three germ layer types. Through much of early embryogenesis, most or all blastomeres can be reprogrammed to adopt a completely different fate when they ectopically express transcription factors that activate the development of very different lineages (Fukushige and Krause, 2005; Gilleard and McGhee, 2001; Horner et al., 1998; Quintin et al., 2001; Zhu et al., 1998). For example, ectopic expression of the GATA transcription factor END-1, which is redundantly required to specify the endoderm founder cell (Maduro et al., 2005), provokes virtually all embryonic cells to adopt an endodermal fate, resulting in an embryo composed entirely of gut (Zhu et al., 1998). In such embryos, essentially all other pathways of differentiation are repressed,

indicating that the normal development of non-endodermal cells is redirected into the gut differentiation pathway. Similarly, ectopic expression of the ectoderm-promoting GATA transcription factors ELT-1 or ELT-3, and the mesodermal factor HLH-1, a homologue of the MyoD family of muscle differentiation factors, causes virtually the entire embryo to develop into skin (Gilleard and McGhee, 2001) and muscle (Fukushige and Krause, 2005) respectively. Further, PHA-4/FoxA can activate pharynx development in many cells normally destined to adopt other fates (Horner et al., 1998). These studies demonstrate that embryonic cells remain pluripotent, i.e., capable of giving rise to cells of all three germ layer types, well after the founder cells are born and specified.

These studies also revealed that the ability of embryonic cells to undergo developmental reprogramming is temporally limited to the early stages of embryogenesis. Specifically, the competency of somatic blastomeres to become reprogrammed into other cell types is rapidly lost starting at the ~100-cell stage (i.e., when the founder cell of the entire endoderm, or E cell, has undergone two rounds of division, henceforth referred to as the 4E stage). By the ~200-cell (8E) stage, all blastomeres appear to have transitioned to a committed state and cannot generally be reprogrammed by transcription factors that direct unrelated developmental pathways. These experiments reveal that a multipotency-to-commitment transition (MCT) occurs during the first half of embryogenesis (Fig. 1). This transition correlates with changes in the compaction of chromatin within the nucleus (Custer et al., 2014; Yuzyuk et al., 2009). A similar event likely occurs in most or all animal embryos (Dang-Nguyen and Torres-Padilla, 2015; Gaspar-Maia et al., 2011; Giammartino and Apostolou, 2016; Hemberger et al., 2009; Koh et al., 2010; Ricci et al., 2017).

The existence of the MCT might indicate that gene regulatory networks appropriate for each cell lineage establish molecular "lock-down" switches (e.g., through positive autoregulation of transcription factors that activate cell-type specific gene expression and differentiation; (Davidson, 2010; Sommermann et al., 2010)) for restricted pathways of differentiation. The lockdown structure of such regulatory networks might resist mobilization of heterologous cell fate-determining gene networks that can be activated by ectopic transcription factor expression prior to the MCT. In addition to the architecture of gene regulatory networks, changes in chromatin organization that preclude transcription factors from accessing their target genes, also restrict developmental plasticity. Studies in *C. elegans* suggest that such processes contribute at least partially to the MCT.

Molecular control of the MCT and temporal extension of cellular plasticity

Mechanisms underlying the MCT have been probed by genetically screening for factors required for its normal onset based on the notion that debilitation of such components would be correlated with altered developmental plasticity (Djabrayan et al., 2012; Yuzyuk et al., 2009). Cells in embryos lacking these factors might be expected to differentiate normally but should be susceptible to reprogramming by ectopic transcription factor expression later than normal.

Based on this logic, the *mes-2* gene was identified by Yuzyuk et al. (2009) as being a global inhibitor of cell fate plasticity during embryonic development. MES-2, the worm homologue of Drosophila Enhancer of zeste, is the catalytic component of a Polycomb repressive complex. The authors observed that forced expression of the muscle-promoting

transcription factor HLH-1, which activates muscle differentiation in ectopic lineages when expressed prior to, but not after, the MCT in 8E embryos, led to a significantly higher proportion of embryos showing widespread expression of a muscle-specific marker and the absence of a pharynx-specific marker in *mes-2* mutants. Transcriptional profiling of 2E, 4E, and 8E stage embryos revealed that *mes-2* mutants fail to downregulate genes whose expression is normally temporally restricted to the 2E and/or 4E stages in wild-type embryos, suggesting that the early phase of gene expression, which is normally attenuated in the period leading up to the MCT, is extended in these mutants. A similar result was obtained in mutants defective for MES-3, which encodes another PRC2 component. A change in the extent of chromatin condensation was inferred to occur around the time of the MCT, based on a decrease in the intranuclear distance between nuclear sites tagged with a fluorescent protein. In *mes-2(-)* mutant embryos, this effect was diminished; i.e., the nuclear fluorescent tags remained more distant. This observation was taken to mean that the PRC2 complex promotes remodeling of chromatin into a state that restricts the accessibility of the genome to cell fate determining transcription factors.

These findings imply that the mechanisms that regulate conversion of embryonic cells from a developmentally plastic to a committed state can be at least partially uncoupled from the events controlling cell fate *per se*. Despite the extension in the period of developmental plasticity, perdurance of transcriptional signatures of early embryogenesis, and a failure to undergo normal compaction, *mes-2* null mutant embryos are viable, and activation of zygotic genes appears to occur normally. These findings imply that commitment does not solely reflect activation of lockdown circuits in gene regulatory networks that control cell identity: developmental events, such as activation of zygotic gene expression, do not specifically

define the transition from multipotency to commitment. Indeed, initiation of cell fate specification does not appear to define the onset of the MCT, since by the 8E-cell stage, many tissue-specific markers are expressed in their respective progenitor cell lineages, and yet cells can be fully transdetermined into several possible cell fates. The H3K27 methyltransferase MES-2 is proposed to globally regulate chromatin reorganization, and perhaps it is this architectural restructuring that determines the onset of the MCT.

The MCT is not exclusively controlled by global mechanisms, such as chromatin remodeling by the PRC2 complex. Rather, in certain lineages, there are additional factors that can determine the point in development in which a lineage becomes committed (Djabrayan et al., 2012). Specifically, Notch signaling, which acts recursively throughout the lineage of the AB founder cell to specify particular cell identities, was found to be required to establish a memory state during development that later determines the timing of onset of the MCT. Using a temperature-sensitive mutant of GLP-1/Notch, Djabrayan et al. (2012) demonstrated that when Notch signaling is eliminated during a critical early period of development, AB-derived blastomeres can be reprogrammed into the endoderm pathway well beyond the normal MCT. In the absence of Notch function, these cells instead maintain developmental plasticity as late as the 20E-cell stage. Unlike previously described roles for Notch in AB lineage specification (Evans et al., 1994; Hutter and Schnabel, 1994; Mello et al., 1994; Moskowitz and Rothman, 1996; Moskowitz et al., 1994; Priess et al., 1987), which involves reception of signals produced by descendants of P_1 (the sister cell of AB), this Notch-mediated restriction of plasticity is autonomous to the AB lineage, implying that both ligand and receptor in this event are expressed by AB descendants. Indeed, knockdown of *dsl-3*, which encodes a non-canonical DSL-like Notch ligand that is predicted to be secreted rather

than membrane-associated, also resulted in AB-autonomous extension in onset of the MCT. Thus, the function of Notch in AB fate specification, and in regulating developmental plasticity in the AB lineage, appear to be mediated by different sets of ligands. Further, *dsl-3* is not essential for embryonic viability, again suggesting that the mechanisms that establish the MCT are separable from those involved in cell fate specification and normal embryonic development.

This action of Notch in promoting differentiation and inhibiting plasticity is not limited to the *C. elegans* embryo: inhibition of Notch was shown to improve the efficiency of iPSC generation from mouse and human keratinocytes, removing the requirement for KLF4 and CMYC, and allowing for reprogramming by only SOX2 and OCT4 (Ichida et al., 2014). Additionally, similar to the AB lineage specificity of Notch in restricting plasticity in *C. elegans*, inhibition of Notch was not found to enhance iPSC generation from mouse embryonic fibroblasts, demonstrating the importance of cellular context in the effect of Notch on reprogramming potential.

Totipotency to commitment: reprogramming in the germline

The germline of *C. elegans* has proven to be a valuable model for investigating stem cell maintenance and regulation (for review see (Joshi et al., 2010)). The germline is organized in a developmentally linear manner. The somatic gonad is composed of two symmetrical U-shaped gonad arms. In the distal-most region of each arm, germ cell progenitors are maintained in a mitotic, proliferative state within a syncytium in response to a Notch signal from a somatic cell that creates the germline stem cell niche (Byrd and Kimble, 2009; Kimble and White, 1981). As nuclei progress proximally and move away from the influence

of this signal, they switch into a meiotic cell division program, arresting at the pachytene stage of Meiosis I, and complete oocyte maturation. As the oocyte possesses the potential to develop into all cells of the animal upon fertilization, there must exist mechanisms that keep this potential in check. Several such mechanisms have been illuminated in *C. elegans* and appear to involve mechanisms that are conserved across animal phylogeny.

A predominant mode of regulation during germline maturation and differentiation in *C. elegans* and other animals is translational control (Merritt et al., 2008; Slaidina and Lehmann, 2014). Many maternal transcripts encoding regulators of embryonic fate specification and differentiation are synthesized during oogenesis. Thus, maintenance of totipotency relies critically on repression of these transcripts during germline development until fertilization and initiation of embryogenesis. The translational regulators GLD-1 and MEX-3 perform complementary functions in maintaining totipotency in the germline (Ciosk et al., 2006). GLD-1, a STAR family of KH domain-containing RNA binding proteins homologous to Quaking and Sam68, regulates the switch from mitosis to meiosis by inhibiting translation of pro-mitotic and anti-differentiation proteins, including MEX-3 (Jones et al., 1996). MEX-3 is expressed in a non-overlapping pattern with GLD-1 in the germline and is predicted to regulate translation of multiple RNA binding proteins, including the putative *C. elegans* homologue of human SOX2, which is required for embryonic stem cell renewal and generation of induced pluripotent stem cells in mammals (Draper et al., 1996; Mootz et al., 2004; Pagano et al., 2009; Pereira et al., 2013). It was found that simultaneous removal of *gld-1* and *mex-3* function causes spontaneous teratoma formation in the meiotic germline, i.e., the premature differentiation of germline cells into differentiated cells from all three germ layer types (Fig. 1). These factors likely function by repressing cell fate-determining

transcription factors and modulating components that regulate chromatin and nuclear architecture, thereby preventing somatic differentiation during germline development.

As with embryos, the totipotency of the germline can be tapped to reprogram cells into somatic cell fates by forced expression of transcription factors (Tursun et al., 2011); such approaches have revealed plasticity-regulating mechanisms that are common to both the germline and the embryo. While overexpression of a single transcription factor alone is insufficient to induce reprogramming of germ cells, ectopic expression of the neuronal selector zinc finger transcription factor CHE-1, in combinations with RNAi knockdown of LIN-53, the worm homologue of the histone-binding protein RbAp46/48, which participates in chromatin remodeling, results in directed conversion of germ cells into ASE-like gustatory neurons (Fig.1) (Tursun et al., 2011). It was subsequently found that LIN-53 acts cell-autonomously through MES-2/MES-3/MES-6, the *C. elegans* Polycomb repressive complex 2 (PRC2), to prevent germ cell reprogramming (Patel et al., 2012). Inhibition of PRC2 also allows germ cells to be reprogrammed into muscle cells when the muscle-determining HLH-1 transcription factor is expressed ectopically. PRC2 therefore appears to function as a global repressor of cellular plasticity both in early embryos (Yuzyuk et al., 2009) and in the germline.

As in the embryo, Notch signaling has been shown to influence the propensity for germ cells to respond to reprogramming cues (Seelk et al., 2016); however, in the opposite direction. Germline nuclei in the stem cell niche are maintained in a proliferative state through the action of Notch signaling from the distal tip cells (DTC) at the end of each gonad arm (Kershner et al., 2013). A gain-of-function mutation affecting GLP-1/Notch causes the germ cells to remain proliferative, resulting in germline tumors. In addition, GLP-1/ Notch

signaling enhances the conversion of germ cells into neurons in response to ectopic CHE-1 expression following knockdown of *lin-53*/RbAp46/48 function (Seelk et al., 2016). It appears that the role of Notch signaling in promoting germ cell conversion into neurons can be uncoupled from its action in promoting germ cell proliferation: reprogramming of germ cells is not enhanced in Notch-independent germ cell proliferation mutants (i.e., in which stem cells divide even without Notch signal (Kadyk and Kimble, 1998; Kimble and Crittenden, 2007)). The separable Notch functions in this germ cell context is reminiscent of the independent roles of Notch in the embryonic AB lineage context, in which it separately promotes cell fate specification and regulates developmental plasticity (Djabrayan et al., 2012). However, it is not clear whether, as in the latter case, there are distinct Notch ligands that mediate these two roles in the germline.

Analysis of differential gene expression in GLP-1/Notch loss-of-function and gain-of-function mutants suggests that Notch permits cell fate conversion by opposing PRC2 action. One identified focal point in this antagonistic relationship is UTX-1, a JmjC domain-containing H3K27me3 demethylase; PRC2 silences *utx-1* expression, while Notch signaling activates its expression (Seelk et al., 2016). *utx-1* transcription is critical for Notch-enhanced germ cell conversion, and therefore UTX-1 may be stimulated to promote reprogramming in the germ cell context.

Studies of germline totipotency in *C. elegans* have uncovered additional regulatory mechanisms of cellular plasticity that appear to be commonly used across phylogeny. The TRIM-NHL (tripartite motif and Ncl-1, HT2A, and Lin-41 domain) protein LIN-41, a novel RBCC (Ring finger B-box and Coiled-Coil) protein that was originally identified as a key heterochronic regulator of larval somatic cell fate (Slack et al., 2000), is essential to prevent

oocytes from prematurely transitioning into embryonic development (Tocchini et al., 2014). Loss of LIN-41 in the germline results in inappropriate activation of the embryonic genome, transition from the meiotic to mitotic cell cycle, and somatic-like differentiation, resulting in teratoma formation. LIN-41 is thought to regulate totipotency in oocytes at the post-transcriptional and/or post-translational level in the cytoplasm, although the precise mechanism is not yet clear (Tocchini et al., 2014). The Drosophila TRIM-NHL homologues Mei-P26 (Neumüller et al., 2008) and Brat (Brain Tumor) (Betschinger et al., 2006) are also specifically associated with cells that have switched from a stem cell state and are undergoing differentiation in the ovary and in asymmetrically dividing neuroblast cells respectively. Loss of Mei-P26 causes the transit-amplifying ovarian cells to proliferate indefinitely, forming a tumor (Neumüller et al., 2008). Brat is asymmetrically segregated to one of the neuroblast cells, where it promotes differentiation and inhibits self-renewal. In the absence of Brat, the transit-amplifying neural progenitor cells continue dividing, forming a tumor (Betschinger et al., 2006; Bowman et al., 2008). Similarly, the murine homologue of these proteins, TRIM32, maintains a balance between self-renewal and differentiation of neural progenitor cells in the neocortex via degradation of proteins such as cMyc by ubiquitination, and an increase in differentiation-promoting miRNA activity, including Let-7 (Schwamborn et al., 2009). The Let-7 miRNA has also been shown to inhibit reprogramming of differentiated cells into pluripotent human iPS cells: inhibition of Let-7 promotes the OCT4-SOX2-KLF4-mediated reprogramming by inhibiting translation of the prodifferentiation factor EGR1 (Worringer et al., 2013). Collectively, these observations suggest that Lin-41/microRNA constitute a conserved pathway that regulates the balance between renewal of pluripotent cells and differentiation.

Violating the MCT: somatic reprogramming of fully differentiated cells by ELT-7

The fully differentiated, post-mitotic somatic cells of *C. elegans* generally maintain a stable cell fate, with the notable exception of two that undergo determinate programmed transdifferentiation (reviewed below). As differentiation of all cells occurs after the MCT, differentiated somatic cells are generally resistant to the type of forced reprogramming that can be observed in developmentally plastic pre-MCT blastomeres in early embryos. While most cells of the newly hatched larva become fully differentiated and post-mitotic prior to the end of embryogenesis, many continue to divide and develop throughout post-embryonic development, long after the MCT. Nonetheless, although these cells and their descendants do not differentiate until later in larval development, they have not been found to be widely plastic. However, particular fully differentiated or developing cells are permissive to forced transdifferentiation in response to specific transcription factors.

The GATA transcription factor ELT-7 has been shown to be capable of overriding the MCT and promoting transdifferentiation of post-mitotic non-intestinal cells into gut-like cells at any stage of development, including in adults (Fig. 1) (Riddle et al., 2013; Riddle et al., 2016). ELT-7 is normally expressed exclusively in the E-cell lineage and throughout the life of the animal in the differentiated gut. It functions as a terminal component of the endoderm gene regulatory network (Maduro, 2017; Sommermann et al., 2010), downstream of the END-1 GATA factor and redundantly with the ELT-2 GATA factor, to initiate gut differentiation during embryogenesis and to maintain transcription of intestinal genes. While ubiquitous expression of various cell fate-determining transcription factors after the MCT has not been reported to initiate target gene expression, a brief pulse of ELT-7 expression at any

stage in *C. elegans* development results in widespread transient activation of an *elt*-2 gene reporter in many cells throughout the animal, indicating that ELT-7 is able to access and activates its targets even in fully differentiated, non-endodermal cells. While this expression gradually fades over the ensuing 48 hours, it persists in cells of two organs -- the proximal somatic gonad, and the muscular feeding organ, or pharynx -- where it is associated with stable transdifferentiation and remodeling of the affected cells. Although the END-1 GATA factor cannot trigger transdifferentiation after the MCT, additional members of the endoderm gene regulatory network, END-3 and ELT-2, also GATA-type transcription factors, are similarly capable of inducing transdifferentiation in *C. elegans*, albeit with somewhat reduced efficacy compared to ELT-7 (Riddle et al., 2016). It is conceivable that the unusually small size of ELT-7 (198 amino acid residues) or other structural characteristics may allow it to access promoters in an otherwise inaccessible chromatin context and thus contribute to its capacity to promote efficient transdifferentiation of differentiated cells.

In the case of the somatic gonad, brief ubiquitous expression of ELT-7 during mid-to-late larval stages redirects the development of the hermaphrodite uterus and spermathecae, and the male vas deferens, all of which arise from equivalent somatic gonadal cell lineages, into a well-formed intestine-like organ (Fig. 2) (Riddle et al., 2016). Within the uterus, an intestine-specific intermediate filament protein, which is normally expressed during terminal differentiation of the embryonic gut (Bossinger et al., 2004), is activated and contributes to the formation of an ectopic intestinal lumen-like structure (Riddle et al., 2016). Electron microscopy of the reprogrammed uterus revealed that it is virtually indistinguishable at the fine ultrastructural level from the normal intestine, including the features of an intestine-like terminal web and microvilli; hence the developing proximal gonad appears to undergo

conversion to a morphologically normal gut in the process of "transorganogenesis." The order and relative timing of these events are similar to those during normal embryonic gut development, suggesting that this process of transorganogenesis involves redeployment of the normal embryonic gut development program. It is of particular note that the *de novo* genesis of an intestine-like organ can occur in a very different context from the normal process: instead of forming within an embryo from naïve, undifferentiated blastomeres, it can be created from a nearly fully developed organ of an entirely different identity. It remains to be determined whether transorganogenesis results in a fully functional gut organ.

The ability to provoke gut development in the developing gonad is limited to the L3-L4 stages of larval development, the period in which the majority of uterine organogenesis, including cellular proliferation, occurs. At the time of hatching, the somatic gonad consists of only two precursor cells, Z1 and Z4, which subsequently divide during larval development into a total of 143 cells, with the uterus alone containing 60 cells (Herman, 2006). Prior to the L3 larval stage, the twelve cells that compose the somatic primordium of the hermaphrodite appear to be insensitive to forced reprogramming. The developmental window for uterus-to-intestine transorganogenesis ceases by the end of the L4 larval stage, just prior to adulthood. Interestingly, in contrast to the uterus, the spermathecae can apparently be reprogrammed into intestine well into adulthood. The cellular context of the developing uterus may be analogous to early blastomeres, which have not yet traversed the MCT. Regardless, these cells are not widely plastic: forced expression of HLH-1 or ELT-1 is unable to reprogram developing uterine cells into muscle or epidermis respectively. Also, as evidenced by the copresence of ectopic ELT-2 protein and EGL-13/SOX domain transcription factor in a

subset of uterine cells, intestine-specific reprogramming can be initiated by ELT-7 in actively differentiating cells (Cinar et al., 2003).

Ectopic ELT-7 is also capable of reprogramming fully differentiated and functional cells in the pharynx (Riddle et al., 2013). Similar to the cells of the somatic gonad, intestine- and pharynx-specific reporter expression, and ultrastructural analysis, indicate loss of pharyngeal cellular identity occurs contemporaneously with acquisition of intestinal cellular identity, suggesting the pharynx undergoes *bona fide* transdifferentiation. However, in contrast to transorganogenesis of the uterus, pharyngeal cells are susceptible to forced transdifferentiation at any larval stage and throughout adulthood (Riddle et al., 2013). It is noteworthy that the pharynx is fully formed, post-mitotic, and becomes functional by the end of embryogenesis. Although it grows larger, it does not undergo further development during larval stages. Hence the conversion of pharynx cells to gut-like cells by a single transcription factor can occur in terminally differentiated, post-mitotic cells that have been functioning for several days.

In contrast to many other examples of transdifferentiation and reprogramming, these observations demonstrate that (a) a single transcription factor is capable of provoking transdifferentiation, (b) transdifferentiation occurs without a requirement for removing any other factors, and (c) it can occur in post-mitotic cells, without a requirement for intercession of a cell division cycle. In addition, this event is associated with wholesale remodeling of cellular ultrastructure.

The specificity of cell fate switching from pharynx-to-intestine and uterus-to-intestine may be attributable to the endogenous expression of another factor that establishes a permissive cellular context in which ELT-7 can direct intestine fate. One candidate for such a

second factor is the *C. elegans* FoxA transcription factor homologue PHA-4, which is known to be expressed in the intestine, pharynx, and developing somatic gonad, and has been extensively studied in the context of its essential role in pharynx organogenesis (Anokye-Danso et al., 2008; Chen and Riddle, 2008; Gaudet and Mango, 2002; Horner et al., 1998; Mango et al., 1994; Zhong et al., 2010). Although knockdown of PHA-4 during larval development was not found to significantly reduce levels of forced reprogramming to intestine of either organ, elimination of PHA-4 during embryonic development, which prevents formation of a differentiated pharynx, was found to block transdifferentiation (Riddle et al., 2016). This result suggests that PHA-4 activity may not be required at the time of ELT-7 induction for cell fate change to occur, but that normal PHA-4-dependent pharynx differentiation is a necessary condition for pharynx transdifferentiation.

Reprogramming of somatic cells into intestinal cells in *C. elegans*, as with any example of cellular reprogramming, raises the question of whether the adopted cell fate is molecularly identical to the normal cells of the same fate, as defined by transcriptional and chromatin state, or if there remains a "memory" or remnant transcriptional activity persisting from the original fate. The reprogramming of somatic cells with a single transcription factor *in vivo* provides a model for investigating the mechanisms of directed cell fate change. How is the expression of a single transcription factor alone capable not only of initiating the endoderm differentiation network within an abnormal cellular context, but also repressing the endogenous cell fate network? Transient coexpression of genes that are never normally expressed in the same cells during development must be reconciled to resolve into a stable cell fate. The activation of the gut differentiation program and simultaneous suppression of the endogenous transcriptional program occurs through a mechanism that is yet unknown. It

is possible that cell fate conversion relies on extensive protein turnover, which removes the factors specific to the initial fate (Fig. 2). During normal embryonic development, proper cell fate decisions often depend on protein degradation through Cullin-RING ubiquitin ligase E3 complexes (Du et al., 2015). Such turnover may be essential for efficient transdifferentiation.

Commitment to unipotency: naturally occurring reprogramming

Cellular reprogramming in *C. elegans* has been shown to occur not only as a result of artificial experimental conditions, but also during normal development in cells that undergo preprogrammed, and highly stereotyped, transdifferentiation. The ability to track single cells continuously throughout *C. elegans* development allows unambiguous assignment of these transdifferentiation events, providing a fruitful model for such *bona fide* plasticity in differentiated cells. In these cases of natural transdifferentiation, fully differentiated and functional epithelial (Y cell) or glial cells (AMso) later convert into a motor neuron (PDA) and a sensory neuron (MCM) respectively (Fig. 1) (Jarriault et al., 2008; Sammut et al., 2015).

One of these two transdifferentiation events has been characterized at the molecular level, providing substantial insights into the process. Long after the MCT, when all cells have committed to a particular cell fate or lineage and widespread cellular plasticity has been lost, the rectal epithelial "Y" cell proceeds through a stepwise process in which it dissociates and migrates away from the rectal tube, undergoes dedifferentiation and loss of its epithelial characteristics, and then redifferentiates into the "PDA" motor neuron, which expresses characteristic neuronal markers and morphology (Richard et al., 2011). As these events have been extensively reviewed, we will briefly summarize these studies here; for additional

reviews see (Becker and Jarriault, 2016; Götz and Jarriault, 2017; Hajduskova et al., 2012; Zuryn et al., 2012).

Studies of the Y→ PDA event have revealed remarkable conservation in mechanisms that regulate cellular plasticity in distantly related metazoans. Through a combination of forward and reverse genetic screens for *C. elegans* mutants that show a persistent Y cell fate, it was found that the initiation of Y→PDA transdifferentiation requires the activity of a chromatin-modifying NODE (Nanog and Oct4-associated Deacetylase)-like complex containing SEM-4/Sall, CEH-6/Oct4, and EGL-27/MTA, along with SOX-2/Sox activity (Kagias et al., 2012). As members of this complex have been shown to be involved in pluripotency of mammalian embryonic stem cells (ESCs) (Heng et al., 2010; Neff et al., 2011; Torres-Padilla and Chambers, 2014), this finding provides a striking link between mechanisms involved in ESC plasticity in mammals and natural transdifferentiation in an animal that has diverged from them since the split of protostomes and deuterostomes ~6-700 Mya. In *C. elegans*, these factors appear to function upstream of the homeodomain-containing transcription factor EGL-1, which is further required for an early step in the Y→PDA identity change (Jarriault et al., 2008; Kagias et al., 2012). The dedifferentiated Y → PDA intermediate exists for only a brief period in development, and is devoid of nearly all epithelial and neuronal markers and morphological characteristics. Nonetheless, this intermediate appears to be unipotent, and its fate, tightly regulated. Challenging the unipotent cell through forced expression of intestine or muscle fate-directing transcription factors fails to redirect its differentiation; transdifferentiation only to the PDA fate has been observed for this cell (Richard et al., 2011). Thus, in this *in vivo* cellular reprogramming context, loss of differentiated cell identity is apparently uncoupled from acquisition of multipotency.

Chromatin factors have also been shown to be intimately involved in other aspects of Y→ PDA transdifferentiation, with their activity regulated at different stages of the process. In early Y→ PDA, H3K4 methylation occurs through the *C. elegans* Set1 complex, which includes *set-2* and *wdr-5.1* among others, and findings suggest that the Set1 complex also interacts with the aforementioned NODE factors (Zuryn et al., 2014). In later stages of Y→ PDA conversion, H3K4 methylation is accompanied by H3K27me3/2 demethylation by JMJD-3.1/KDM6B, which can physically interact with both WDR-5.1, and UNC-3/EBF, a transcription factor required for redifferentiation to a neuronal fate (Richard et al., 2011; Zuryn et al., 2014).

The ability of the Y rectal epithelial cell to undergo transdifferentiation is established when the cell is born during embryogenesis and appears to be triggered by LIN-12/Notch signaling (Jarriault et al., 2008). In constitutively active *lin-12* mutants, the sister of Y also adopts a Y cell fate, and later in development animals with two PDA neurons are observed, indicating that both Y cells have undergone transdifferentiation. By using a temperature-sensitive *lin-12* mutant, Jarriault et al. (2008) demonstrated that if Notch function is eliminated after Y cell fate has been established, but before initiation of transdifferentiation, a PDA cell can still arise. Thus, while Notch is required for induction of Y cell fate, it is dispensable for the transition to neuronal cell fate. This Notch functionality may be analogous to the role of Notch signaling in germ cell conversion by antagonizing the repressive heterochromatin state established by PRC2 through expression of histone demethylases (Seelk et al., 2016). Notch signaling may similarly establish a "memory state" in the Y cell which later permits the Y cell to lose its differentiated identity and gain plasticity (unipotential in this case).

A second example of natural transdifferentiation in *C. elegans* involves the sex-specific conversion of fully differentiated and functional glial cells, the bilateral AMso cells, into male-specific neurons, the MCMs, which are involved in sexual conditioning behavior (Sammut et al., 2015). In hermaphrodites, the AMso cells are terminally differentiated and do not undergo this event; plasticity of these cells is triggered during late larval development by the sex-determination pathway. The AMso→ MCM transdifferentiation event differs from that of the post-mitotic Y→ PDA event in that a round of cell division precedes acquisition of the new cellular identity. The molecular events underlying AMso→ MCM transdifferentiation have not yet been elucidated and it will be of interest to learn whether similar mechanisms to those directing transdifferentiation of Y are involved.

Conclusions and Perspectives

A diversity of cellular reprogramming strategies are now known in *C. elegans*. Although no universal mechanisms have emerged, common themes are evident. It is likely that in some cases the *in vivo* cellular context provides a reprogramming-competent state that permits cell fate conversion to a particular end. In mice, induction of murine cardiomyocyte-like cells by Gata4, Mef2c, and Tbx5 overexpression is inefficient *in vitro* but much more robust *in vivo* (Ieda et al., 2010; Qian et al., 2012)**.** This type of cellular environment-enhanced reprogramming in *C. elegans* could be involved in, for example, transorganogenesis of the uterus into intestine. However, cell-autonomous mechanisms act in other settings, such as PRC2-mediated inhibition of germ cell reprogramming (Patel et al., 2012).

It is not unexpected that epigenetic regulation has been implicated as a key mechanism in the MCT and in nearly all reprogramming events in *C. elegans*, often functioning as an

"epigenetic barrier" (e.g., (Watanabe et al., 2013)). In blastomere reprogramming, cells are initially intrinsically plastic, and heterochromatin is not readily observed. PRC2 activity initiates commitment and loss of plasticity, and is a major contributor to the MCT. In germline reprogramming, totipotency is maintained largely through post-transcriptional and post-translational regulation (Lee and Schedl, 2006; Merritt et al., 2008), although a role is also played by epigenetic silencing mechanisms which, if inhibited, can allow directed acquisition of cell fate. Natural transdifferentiation involves stepwise action of epigenetic modifications, and while not absolutely required for successful transdifferentiation, they ensure robust cell fate conversion. By contrast, for somatic reprogramming initiated by ELT-7, removal of inhibitory chromatin marks is remarkably not a necessary step. The specificity of ELT-7-mediated reprogramming into intestine could, however, be aided by an intrinsic absence of repressive chromatin, recruitment of endogenous chromatin modifying complexes, or early expression of a tissue-specific chromatin remodeller. With respect to the last possibility, transdifferentiation of mesoderm in mouse embryos into cardiomyocytes is achievable by ectopic expression of Gata4, Tbx5, and Baf60c - a cardiac-specific subunit of BAF chromatin remodelling complexes, which may permit binding of Gata4 to cardiac genes in unrelated cell types (Takeuchi and Bruneau, 2009).

Regarding the ability of ELT-7 to override the MCT, it is unclear how a single transcription factor is capable of inducing both transdifferentiation of the pharynx, and transorganogenesis of the somatic gonad, or why the latter is temporally restricted while the former is not. Although evidence suggests that deployment of intestinal differentiation follows a sequence similar to that in the embryo, the mechanism that promotes establishment of an ectopic cellular identity could be shared, or perhaps entirely distinct, depending on the

cellular context. Also unclear is the response and ultimate fate of other cells and tissues that ectopically express transcriptionally active ELT-7 and an early readout of gut differentiation, yet do not show signs of a stable change in cell fate. Additional ectopic ELT-7 expression phenotypes include developmental arrest, impaired chemotaxis, and molting defects, suggesting that there may be more widespread changes in gene expression that disrupts normal functioning of epidermis and neurons, for example, but is insufficient for full transdifferentiation into intestine. This could be the result of additional repressive mechanisms that are not present in the pharynx or somatic gonad. If so, then removal of such factors might make it possible to reprogram many more, or even all cells of the animal.

A particularly unusual characteristic of reprogramming identified in *C. elegans* is the requirement for the establishment of a fully differentiated cellular identity for successful transdifferentiation. This appears to be the case for both natural transdifferentiation, in which Notch-dependent specification of rectal epithelial cell fate is necessary for later conversion to neuronal cell fate (Jarriault et al., 2008), and for forced transdifferentiation, in which PHA-4-dependent specification of pharyngeal cell fate is necessary for subsequent ELT-7-induced transdifferentiation to an intestinal cell fate (Riddle et al., 2016). Intriguingly, neither Notch nor PHA-4 appears to be required at the time of cell fate conversion. More likely these factors act to establish a permissive cellular context that allows endogenous or ectopic factors to initiate transdifferentiation. Interestingly, PHA-4/FoxA exemplifies a class of transcription factors called Pioneer factors that have been shown to bind to target sequences on nucleosomes associated with “silent” chromatin (Cirillo et al., 2002; Fakhouri et al., 2010) and can initiate the establishment of a chromatin state that is competent to be activated (Zaret and Mango, 2016). Similarly, the iPSC-promoting transcription factors Oct4 and Sox2,

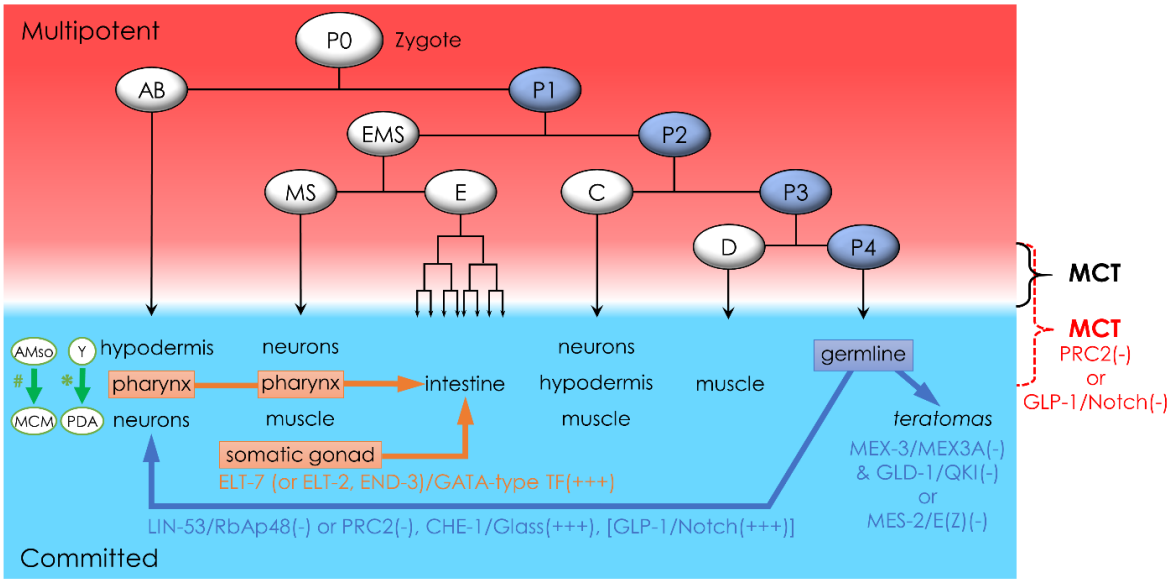
which are also required for Y→ PDA transdifferentiation, behave as Pioneer factors by binding to transcriptionally silent chromatin sites in human fibroblasts prior to their transformation into iPSCs (Soufi et al., 2012).

Revealing the mechanisms that regulate cell fate plasticity in *C. elegans* has the potential to enhance our understanding of human disease. For example, some cell types appear to be more prone to conversion in fate which, in humans, can lead to metaplasias, the appearance of inappropriate differentiated tissue at ectopic sites (Slack, 2007; Giroux and Rustgi, 2017). Indeed, Barrett's metaplasia, a precancerous condition linked to esophageal adenocarcinoma, in which the distal esophagus (foregut) is converted to intestine-like epithelium (midgut) (Slack, 2007; Colleypriest et al., 2010; Slack et al., 2010) is analogous, and perhaps homologous, to the pharynx (foregut) → intestine (midgut) transformation promoted by ELT-7 in *C. elegans*. Thus, the pharynx → intestine transdifferentiation event might provide a useful model for understanding this metaplasia, particularly in light of the fact that the mechanisms that distinguish the major divisions (foregut, midgut, and hindgut) of the digestive tract appear to be largely conserved across metazoan phylogeny (Stainier, 2005).

Elucidating the mechanisms of reprogramming in *C. elegans* and their connections with the MCT may provide insights into conditions that involve inappropriate loss or gain of plasticity, resulting in impaired healing after injury, or oncogenesis. Studying cellular reprogramming in a model organism such as *C. elegans* also benefits from the ability to investigate the phenomena in an *in vivo* cellular context. *In vivo* reprogramming of cells to repair damaged or lost tissues has now been proposed as a potential treatment for various conditions, including ischemic injury, and neurodegeneration (Banga et al., 2012; Heinrich et al., 2015; Li and Chen, 2016; Li et al., 2014; Niu et al., 2013; Qian et al., 2012; Taguchi and

Yamada, 2017; Srivastava and DeWitt, 2016; Zhou et al., 2008). Thus, more efficient and robust methods for converting tissues for regenerative medicine could greatly benefit from mechanistic understanding of cell fate commitment and lineage reprogramming in *C. elegans* and other model systems.

Fig 1.



Natural Transdifferentiation: * *lin-12*/Notch, *sem-4*/Sal, *egl-5*/Abd-B, *unc-3*/EBF, *egl-27*/MTA1, *ceh-6*/POU3F4, *sox-2*/SOX, *jmjd-3.1*/JMJD3, *set-2*/Set1

sex-determination pathway, cell division

Fig 1. Cellular context of reprogramming in *C. elegans*.

Early cell divisions in *C. elegans* embryonic development establishes six founder cells which give rise to specific differentiated tissue types. Despite early lineage specification, all somatic embryonic blastomeres are multipotent (*red background*) until approximately the stage at which the E-lineage has generated eight cells (8E-cell stage). A major developmental transition, the Multipotency-to-Commitment Transition (MCT), then occurs, after which blastomeres are committed to their respective specified fates (*blue background*). Multipotency can be extended temporally, resulting in a delayed MCT with respect to developmental progress, through inhibition of PRC2 or Notch signaling (*red text*). Post-MCT cell fates can be converted through natural or forced transdifferentiation (*arrows*). Inhibition of MEX-3 and GLD-1, or inhibition of MES-2 causes germ cells to spontaneously differentiate, resulting in teratomas. Inhibition of LIN-53 or PRC2, in combination with forced expression of CHE-1, causes germ cells to be reprogrammed into neurons, and conversion can be enhanced by increasing Notch signaling (*blue text*). Forced expression of ELT-7 (or ELT-2, or END-3) causes transdifferentiation of the fully differentiated and functional pharynx, and transorganogenesis of the developing somatic gonad, into intestine, with the uterus forming a secondary gut-like organ (*orange text*). During larval development, a single epithelial cell ("Y") undergoes programmed transdifferentiation into a neuronal cell ("PDA") through a natural process involving numerous identified factors; during male-specific larval development the glial amphid socket cells ("AMso") undergo an additional cell division, in which one daughter self-renews as a glial cell while the other loses glial characteristics and acquires neuronal characteristics, and transdifferentiates into a male-

specific neuron (“MCM”) (*green text*). Loss of gene function is indicated by (-) and gain of function or forced expression is indicated by (+++).

Fig 2.

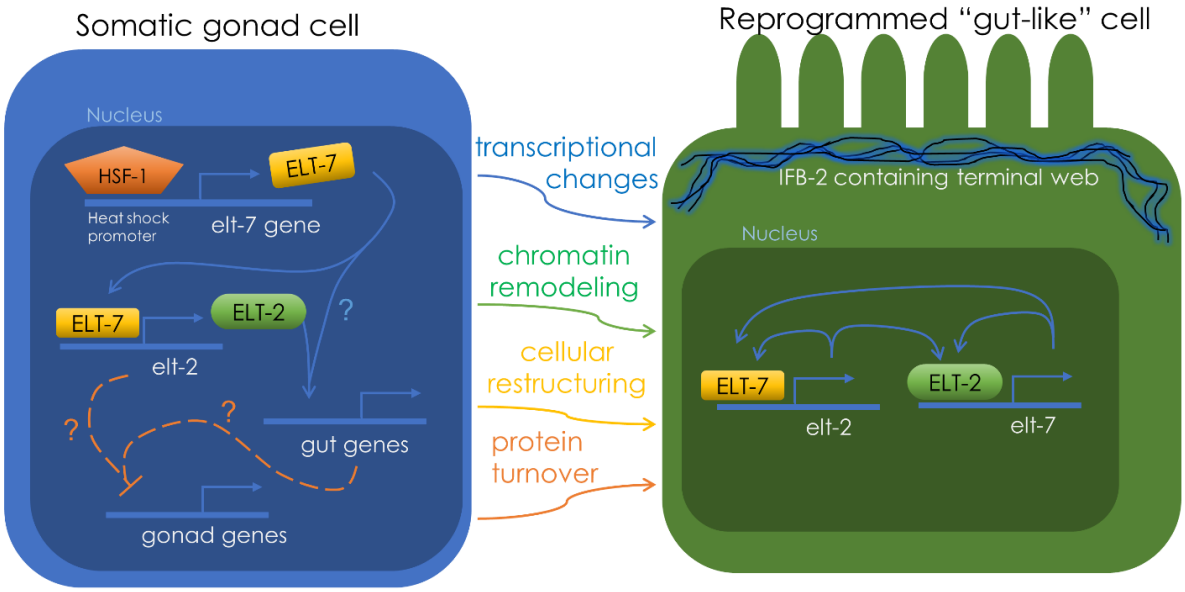

Somatic gonad cell
Reprogrammed "gut-like" cell
Nucleus
HSF-1
ELT-7
Heat shock promoter
elt-7 gene
ELT-7
ELT-2
elt-2
?
?
?
gut genes
gonad genes
transcriptional changes
chromatin remodeling
cellular restructuring
protein turnover
IFB-2 containing terminal web
Nucleus
ELT-7
ELT-2
elt-2
elt-7

Fig 2. Events in somatic gonad → intestine transorganogenesis.

Forced expression of the ELT-7 transcription factor activates a positive feedback loop that "locks-down" the gut differentiation pathway, resulting in activation of downstream gut genes in the pharynx and somatic gonad and repression of genes involved in the original differentiated state. The proximal somatic gonad undergoes transorganogenesis, during which the reprogrammed "gut-like" cells are converted to those with ultrastructural characteristics indistinguishable from that of the endogenous intestine, an event that requires changes in gene expression, protein turnover, and cellular remodeling.

Chapter Two

Probing the potency of the transcription factor ELT-7 to force direct cell fate reprogramming

Abstract

Widespread ectopic activation of the ELT-7 transcription factor is capable of forcing direct cell fate conversion in fully differentiated and post-mitotic tissues. The mechanisms underlying this potent reprogramming ability, however, remain poorly understood. To explore the role of ELT-7 in reprogramming and to search for additional factors that can enable or prevent transdifferentiation into intestine-like tissue, multiple genetic approaches were utilized, including quantification of transgenic intestine-specific reporter constructs, *elt-7* deletion mutants, and targeted RNAi screening. I find that *elt-2p::gfp* and *ifb-2::cfp* reporter expression follows a reproducible and tissue-specific pattern during reprogramming, and that END-3-driven reprogramming is successful in an *elt-7(-)* mutant. I identify three chromatin-associated factors, *epc-1, smo-1,* and *pyp-1* that may be important for somatic gonad-to-intestine transorganogenesis. These results suggest that direct reprogramming to an intestine cell fate involves recapitulation of embryonic-like intestine developmental events. In addition, the complete formation of a secondary intestine in worms without a functional *elt-7* gene shows that ELT-7 activity is not necessary for transorganogenesis. The vast majority of factors tested in the RNAi screens had no post-embryonic RNAi phenotypes or effects on reprogramming efficiency. In contrast, the three chromatin-associated factors identified all have an effect on gonad development and their inhibition interferes with ELT-7 directed reprogramming. These factors may each function in the establishment of a chromatin state that is permissive to reprogramming to an intestine fate, although the molecular mechanisms involved remain to be elucidated. These studies provide insight into the dynamics of *in vivo* reprogramming, affirm the robust capability of the endoderm gene

regulatory network to redirect cell fate, and highlight factors that may determine the cellular contexts that allow for reprogramming into intestine in *C. elegans*.

Introduction

Embryonic development of the *C. elegans* intestine involves a core pathway of seven transcription factors that are activated in a sequential feed-forward cascade (Fig 1 and see also Chapter 1) (Maduro, 2015; Maduro et al., 2007; Maduro et al., 2015; Sommermann et al., 2010). This regulatory logic enables rapid and robust specification of intestine cell fate (Maduro, 2009; Maduro, 2015; Maduro et al., 2005; Raj et al., 2010). The rest of the cell types in the embryo are similarly specified and early in development the embryo completes the multipotency-to-commitment transition (MCT), after which cell fates stabilize and remain fixed and resistant to alteration for the remainder of the worm's life (Djabrayan et al., 2012; Erdelyi et al., 2017; Rothman and Jarriault, 2019; Spickard et al., 2018; Yuzyuk et al., 2009). Previous research has revealed a novel and remarkable exception to this norm; multiple factors within the *C. elegans* endoderm developmental transcription factor cascade, namely END-3, ELT-2, and ELT-7, each individually possess the ability to force cell fate changes resulting in the complete restructuring of an entire organ within the living animal (Riddle et al., 2013; Riddle et al., 2016). It is the ELT-7 transcription factor, however, that induces cellular reprogramming with the greatest efficiency.

What makes this particular factor such a potent activator of intestine gene expression in an ectopic cellular context? If ELT-7 contains a unique reprogramming ability, the interconnectivity of the endoderm gene regulatory network (GRN) could function to activate ELT-7 from various initially expressed nodes. Alternatively, the robustness of the network as a whole may drive intestine fate acquisition regardless of the initial starting point. In addition to ELT-7 activity, what are the molecular characteristics that define the particular cellular contexts in which reprogramming into intestine is permissible?

In order to address these questions, I used RNA interference and genetic mutants to perturb the endoderm GRN and screened for factors that modulate the reprogramming ability of ELT-7. I also quantified the ectopic expression of intestine tissue-specific reporter constructs to characterize the dynamics and sequence of *in vivo* reprogramming events. I tested 338 genes encoding nuclear proteins known to act directly on DNA or chromatin and performed analysis of intestine reporter expression to identify subtle and quantitative differences in tissue-specific reprogramming, as well as rare, major effects on ELT-7 directed transorganogenesis efficiency.

In this chapter I show that transdifferentiation of the pharynx and transorganogenesis of the somatic gonad into intestine both follow a sequence of events similar to embryonic intestine development, but with differing temporal dynamics, suggesting overlapping but partially distinct reprogramming mechanisms in these tissues. I also show that somatic gonad-to-intestine transorganogenesis can occur in worms without functional ELT-7, revealing that this factor is not necessary for reprogramming to an intestine fate. I further report the results of the candidate nuclear factor RNAi screening. Although the majority of gene knockdowns had no effect on ELT-7 directed reprogramming, I found three factors that appear to play a role in transorganogenesis: *epc-1*/E(Pc), *smo-1*/SUMO1, and *pyp-1*/PPA1. These studies affirm the potency of ELT-7 and the intestine GRN to manipulate cell fates and provide insight into the mechanisms of *in vivo* cellular reprogramming and transorganogenesis.

Results

Gut specific transgene markers suggest redeployment of gut development in pharynx and somatic gonad tissues

The dramatic changes in cell fate after widespread activation of the transcription factor ELT-7 in *C. elegans* have been confirmed by multiple lines of evidence when observing worms around 48 hours post activation (Riddle et al., 2013; Riddle et al., 2016). Most striking is electron microscopic observations revealing the reprogrammed uterus to be virtually indistinguishable from the endogenous intestine at the ultrastructural level (Riddle et al., 2016). While this endpoint of transorganogenesis is remarkable, the dynamics of the process are not well characterized. The rate and efficiency of reprogramming in the cells of the pharynx and somatic gonad could indicate how cells respond to ectopic ELT-7. Cells that are more robust in their GRNs and that resist reprogramming more or less strongly may have different dynamics from cells that are more susceptible to reprogramming. Furthermore, as there are different tissues that become reprogrammed, there could be shared or disparate mechanisms that enable them to be reprogrammed into intestine-like cells. Similar temporal dynamics might suggest a similar mechanism, while different dynamics might suggest different mechanisms.

The dynamics of cellular reprogramming were determined by quantifying the percent of worms in a population that expressed either *elt-2p::gfp* or *ifb-2::cfp* transgenic markers in a particular tissue over a range of time points. The *elt-2p::gfp* transgene is a transcriptional reporter expressed early in the E-cell lineage and is continually expressed into adulthood. The *elt-2* gene can be activated by ELT-7 (Sommermann et al., 2010), so the transgene also

serves as a marker for ELT-7 transcriptional activator activity and an early marker for potential conversion to gut cell fate. The *ifb-2::cfp* transgene is a translational reporter specifically expressed in the intestine from late embryonic development throughout adulthood (Bossinger et al., 2004; Hüsken et al., 2008). *ifb-2* encodes an intermediate filament protein which is integrated into the terminal web cytoskeleton of the intestine and is an indicator of differentiation into intestinal cell fate (Dodemont et al., 1994; Jahnel et al., 2016; Karabinos et al., 2001). The experiment was carried out by feeding multiple plates of synchronized *hsp::elt-7* transgenic L1 larvae for approximately two days, until the early L4 stage, and then exposing the worms to a 33^0C heat shock for 30 minutes. For each time point a separate population of worms was scored for whether *elt-2p::gfp* and *ifb-2::cfp* was prominently expressed in either the pharynx, somatic gonad, or other tissues (Fig 2).

After widespread activation of *hsp::elt-7*, *elt-2p::gfp* is rapidly expressed and by 6 hours GFP is visible in all major cell types, including hypodermis, muscles, and neurons. It is particularly strongly expressed in the pharynx (Fig 2A-D). *elt-2p::gfp* expression is more gradually observed in the somatic gonad and on occasion is absent in worms that later contain a well-formed ectopic lumen. The expression of *ifb-2::cfp* takes longer to be visibly expressed than *elt-2p::gfp* in all tissues. It takes 8-12 hours on average for *ifb-2::cfp* to be detectably expressed in the pharynx, and 12-16 hours in the somatic gonad (Fig 2G). Ectopic *ifb-2::cfp* is also expressed at very low levels in various tissues other than the intestine, pharynx, and somatic gonad (Fig 2A). Expression of "widespread" *ifb-2::cfp* increases modestly over time until the point when worms are very sick or necrotic. This widespread expression is in the form of small puncta, appearing individually or in small clusters scattered throughout the animal, often in the hypodermis. Owing to their faint and stochastic nature,

precise quantification of the size, intensity, or number of widespread ectopic *ifb-2::cfp* puncta has not been achieved.

In embryonic development, *elt-2* is activated early in the endodermal blastomeres (Fukushige et al., 1998), and *ifb-2* is a downstream target that is activated a few hours later during differentiation of the intestine (McGhee et al., 2009). The dynamics of *elt-2p::gfp* and *ifb-2::cfp* expression after widespread ELT-7 activation follows the same temporal order, which suggests that transdifferentiation of pharynx and/or somatic gonad cells into intestine-like cells may occur similarly to embryonic differentiation of the intestine. There is, however, a notable difference in the dynamics of *elt-2p::gfp* and *ifb-2::cfp* expression in the pharynx versus the somatic gonad. This raises the possibility that there is a different mechanism for each tissue that enables ELT-7-directed reprogramming. Other evidence exists to support this: for example, while PHA-4/FoxO activity during embryogenesis is necessary for later reprogramming, knockdown of *pha-4* during larval somatic gonad development does not prevent reprogramming of the uterus and spermatheca (Riddle et al., 2016).

Widespread ectopic *elt-2p::gfp* expression in other tissues is transient compared to that in pharynx and somatic gonad but is still visible for over 48 hours in most worms. Although widespread ectopic *ifb-2::cfp* expression is relatively insignificant, indicating that these other tissues are not being fully reprogrammed, the prolonged duration of *elt-2p::gfp* could suggest that in a few other cell types ELT-7 continues to activate transcription of target genes. Continued ELT-7 activity could significantly disrupt the transcriptional state of widespread tissues as they would need to actively repress intestine gene networks while maintaining their correct regulatory network. This idea is explored further in Chapter 3. Alternatively, GFP is a

fairly stable protein (Bokman and Ward, 1981; Chalfie, 1995), and a slow fading away of a briefly expressed reporter could be explained by a long half-life.

ELT-7 is sufficient, but not necessary, for cellular reprogramming

Widespread ectopic activation of a single transcription factor is capable of inducing somatic gonad-to-intestine transorganogenesis. The transcription factor that achieves this with the highest efficiency is ELT-7; however, both ELT-2 and END-3 are also able to induce transorganogenesis (Riddle et al., 2016). All of these proteins are GATA factors that are specifically expressed in the E-cell lineage to promote intestinal cell fate (Fig 1). ELT-7 and ELT-2 are both terminal differentiation factors that activate transcription of each other and themselves. END-3 acts more upstream and is transiently expressed during embryonic development, and it has also been observed to have the lowest reprogramming efficiency of the three GATA factors.

Does the ability to induce transorganogenesis depend on the ability to activate ELT-7 in a given cell type? The reprogramming pathway might conceivably converge on one particular transcription factor in the endoderm GRN. There could be a particular biochemical property of ELT-7 that enables it to initiate cell fate change; for example, the ELT-7 protein is smaller than ELT-2, the larger ELT-7 isoform contains 198 amino acids and has a molecular weight of 23.1 kD which may allow it to penetrate deeper into condensed chromatin in order to activate intestinal genes. ELT-2, on the other hand, contains 433 amino acids and has a molecular weight of 47.1 kD. If such is the case, reprogramming initiated by ectopic expression of ELT-2 or END-3 would critically depend on being able to activate the endogenous *elt-7* gene in additional tissues.

To test whether *elt-7* is necessary for successful transorganogenesis I crossed the *elt-7(ok835)* mutation into the *hsp::end-3* strain. The *ok835* mutation consists of a 498 bp deletion which affects the 3' intron and coding exon of the *elt-7* gene (C. elegans Consortium, 2012). This deletion mutation is assumed to result in a loss of function, although *elt-7* is not an essential gene and *elt-7(-)* worms appear wild type. For reasons that are not obvious, the growth rate of the *hsp::end-3; elt-7(-)* strain was slightly slower than the *hsp::end-3* strain. Synchronized L4 worms from both the *hsp::end-3* control strain and *hsp::end-3; elt-7(-)* strain were heat shocked and scored for reprogramming after 24 hours (Fig 3). In the *hsp::end-3* strain, 76% (n = 50) of worms show some signs of somatic gonad-to-intestine reprogramming as determined by the presence and morphology of ectopic IFB-2::CFP. In the *hsp::end-3; elt-7(-)* strain, 57% (n = 37) of worms show signs of reprogramming. Although the loss of *elt-7* results in a decrease in overall *hsp::end-3* reprogramming efficiency, at a $p < 0.05$ significance level, this difference is not statistically significant (Fisher's exact test p-value = 0.067).

These results indicate that an *elt-7(-)* mutation does not prevent somatic gonad-to-intestine transorganogenesis driven by ectopic *end-3* expression. Therefore, although ELT-7 is sufficient to induce transorganogenesis, ELT-7 transcription factor activity is not necessary for this process to occur.

Targeted reverse genetic screening for factors that determine susceptibility to reprogramming by ELT-7

Post embryonic, differentiated cells generally have very stable fates throughout an organism's life span – an essential feature for continued health and longevity. In order to

reprogram cells, either from one fate directly to another or from a differentiated to a pluripotent state, manipulation of multiple factors is often required (Cahan et al., 2014; Firas and Polo, 2017; Gao et al., 2017; Morris et al., 2014). In the now classic example of reprogramming mammalian fibroblasts into induced pluripotent stem cells (iPSCs), four transcription factors were required, and the process was still highly inefficient (Takahashi and Yamanaka, 2006). Numerous subsequent studies that improved iPSC generation efficiency involved manipulation of additional factors, often inhibiting their function (Takahashi and Yamanaka, 2016; Vierbuchen and Wernig, 2012).

Only in rare instances can the activity of a single factor cause the conversion from one cell fate to another (Chambers and Studer, 2011; Chanda et al., 2014; Halder et al., 1995; Weintraub et al., 1989). And yet, in *C. elegans*, activation of the *elt-7* gene alone is capable of reprogramming the pharynx and somatic gonad into intestine-like tissue (Fig 4). On the other hand, the ectopic expression of *elt-7* in all cells of the worm appears to be insufficient for reprogramming tissues other than the pharynx and somatic gonad. What, then, are the molecular characteristics that can define a cellular context that is permissible to reprogramming to an intestine fate?

I performed a series of RNAi screens to identify candidate factors that are involved in determining susceptibility to ELT-7 directed reprogramming. The screens were designed to simultaneously test two complementary hypotheses. First, I hypothesized that there is a particular gene that enables ELT-7 to reprogram the pharynx and/or somatic gonad, the knockdown of which would result in a decrease in the efficiency of reprogramming. Second, I hypothesized that there is a gene that is expressed in certain tissues that functions to prevent reprogramming by ELT-7. Knockdown of such a gene would result in the reprogramming of

one or more additional tissue types. In theory, nearly any gene in the genome could play a role in reprogramming susceptibility. However, maintenance and manipulation of cell fate must depend on the cell's overall transcriptional state. Therefore, factors acting within the nucleus, and especially on chromatin and DNA directly, were determined to be the top candidates for RNAi screening. Two separate methods were used to generate gene lists for RNAi screening and are described below. Furthermore, owing to the labor-intensive requirements for scoring reprogramming phenotypes, it was feasible to screen only ~2% of the genome, or about 350 genes using this approach.

Transcription factor binding site (TFBS) enrichment informed RNAi screening

The primary objective of this screen was to identify genes that enable somatic gonad-to-gut transorganogenesis. A pilot transcriptional profiling experiment using mRNA-seq detected thousands of transcripts that are differentially expressed 3 hours after ectopic *elt-7* expression [see Chapter 3 for discussion of the full transcriptional profiling study]. These early gene expression changes could be attributable to direct transcription factor activity of ELT-7 but may also be driven by increased activity of another, unknown transcription factor, that is promoting cellular reprogramming. In order to identify candidate transcription factors that contribute to early gene expression changes in worms undergoing transorganogenesis, I used the web-based tool oPOSSUM (http://opossum.cisreg.ca/oPOSSUM3/), which enables the detection of over-represented conserved TFBS in sets of genes (Kwon et al., 2012; Sui et al., 2005; Sui et al., 2007). Multiple oPOSSUM analyses were performed using the 2,000 most significantly upregulated transcripts at 3 hours post heat shock (hrs PHS) and TFBS enrichment scores were averaged to determine the transcription factor families with the most prevalent binding sites (Fig 5). Binding sites for GATA factors were the most significantly

enriched, followed by Forkhead, Homeobox, and High Mobility Group (HMG) transcription factors. To populate a candidate list of transcription factors for RNAi screening, most members of the top families, as well as *C. elegans* homologs of the specific transcription factors with the most significant enrichment scores in the oPOSSUM database, were included. For example, all GATA factors in the *C. elegans* genome were screened, plus factors such as the vertebrate beta-beta-alpha-zinc finger transcription factor Evi1 homolog *egl-43*. Additional candidate factors of interest, such as the *C. elegans* chromodomain-containing gene *cec-4*, were included for a total of 118 genes tested in this screen; the L4440 empty RNAi vector was used as the control (Fig 6). Specific results are described in following sections.

Chromatin-related factor RNAi screening

Forced cell fate reprogramming likely involves extensive chromatin remodeling and restructuring of DNA topology to allow for the required dramatic changes in gene expression (Beagan et al., 2016; Guo and Morris, 2017; Onder et al., 2012). Chromatin modifying factors that are specifically expressed in the pharynx or somatic gonad, either endogenously or in response to ectopic ELT-7, could be promoting reprogramming. For instance, a histone acetyl transferase complex that is active in the cells of the uterus may help to establish a chromatin state that is more permissible to ELT-7 binding, the knockdown of which would result in decreased transorganogenesis efficiency. Alternatively, chromatin-related factors expressed in other tissues might establish epigenetic barriers to ELT-7 directed reprogramming, such as H3K9 methyltransferases, which can define transcriptionally silent heterochromatin (Ahringer and Gasser, 2018). Inhibition of such factors might allow for additional tissues to be reprogrammed by ELT-7 into intestine.

A previous screen for chromatin-related epigenetic barriers to cellular reprogramming of germ cells into neuronal cells in *C. elegans* successfully identified the histone chaperone LIN-53/RBBP4 (Tursun et al., 2011). To generate a candidate list of chromatin related genes for RNAi screening I combined lists of *C. elegans* chromatin factors from this previous study and other sources (Cui and Han, 2007; Tursun et al., 2011; Zuryn et al., 2014), and selected 220 genes to be targeted by RNAi (Fig 7).

The majority of gene targets showed no discernable developmental or reprogramming phenotypes when knocked down by RNAi. Knockdowns that had an effect on reprogramming efficiency typically also had more severe developmental phenotypes such as a protruding vulva or sterility. In ELT-7 reprogrammed, L4440 RNAi controls, 89.4% (n = 772) of worms were scored as having an ectopic lumen ("lumen"). The most significant hits were considered to be conditions where ≤ 60% of worms showed the "lumen" reprogramming phenotype, resulting in 19 potential hits. It should be noted, however, that the "wide lumen," and spermatheca reprogramming ("spth") phenotypes are typically indicative of worms that were at a slightly later stage of development when heat shocked and does not necessarily indicate a defect in reprogramming. Follow-up experiments on the most striking or unusual results are described in the next sections.

Knockdown of *epc-1* inhibits transorganogenesis but can provoke heat shock-independent ectopic *elt-2* expression

Initial RNAi screening results identified *epc-1(RNAi)* as having a significant inhibitory effect on reprogramming (Fig 6). However, *epc-1(RNAi)* also shows a strong phenotype in non-heat shocked worms, causing slow growth, developmental arrest and lethality (Fig 9A). The *epc-1* gene encodes Enhancer of polycomb 1, possessing histone acetyltransferase

(HAT) activity, and is a putative member of a conserved Tip60/NuA4-like HAT complex (van Attikum and Gasser, 2005; Ceol and Horvitz, 2004). In nearly all RNAi conditions tested, reprogramming of the pharynx as indicated by ectopic *elt-2p::gfp* and *ifb-2::cfp* was observed in > 95% of worms 24 hours after heat shock with accompanying reprogramming of the somatic gonad in ~90% of worms. For the RNAi screening results, reprogramming of the pharynx is implied in conjunction with somatic gonad reprogramming phenotypes. However, *epc-1(RNAi)* appears to decouple pharynx and somatic gonad reprogramming; therefore, in subsequent experiments, paired phenotypic scoring was used for ectopic lumen formation with or without ectopic gut expression in the pharynx (Fig 8).

Compared to 89% (n = 27) of control worms that have ectopic gut expression in both the pharynx and somatic gonad, for *epc-1(RNAi)* 58% (n = 86) of worms had ectopic gut expression in the pharynx (Fig 9C), and only 10% had both pharynx and somatic gonad ectopic expression. Furthermore, 42% of *epc-1(RNAi)* treated worms did not show signs of reprogramming in the pharynx following ectopic *elt-7* expression, although almost half of these showed signs of reprogramming in the somatic gonad. In control worms that are not heat shocked, no ectopic gut expression is observed (0%, $n \geq 10$) (Fig 8). Surprisingly, in non-heat shocked *epc-1(RNAi)* worms, 20% (n = 45) showed stochastic ectopic gut expression in either the pharynx or somatic gonad, although expression in both was not observed (Fig 9B).

In the standard RNAi screening protocol worms were allowed to feed on RNAi bacteria for 42-45 hours and then heat shocked at the early L4 stage; however, *epc-1(RNAi)* worms phenotypically more closely resembled L2 or L3 stage worms at the time of heat shock in regards to size and vulva/gonad development. In order to mitigate the early larval arrest,

synchronized L1 worms were fed on OP50 bacteria for ~24 hours and then transferred to *epc-1(RNAi)* for an additional day. Worms made more progress developmentally and had L4-like gonads, although were still roughly L3 in size (Fig 9D). When these worms were heat shocked and scored after 24 hours, they showed ectopic IFB-2::CFP patterns that more closely resembled controls (Fig 9E). These results suggest that *epc-1* may not be essential for transorganogenesis but rather delays development and the time window in which somatic gonad reprogramming is permissible. Additionally, as *epc-1* is part of a conserved Tip60/NuA4 complex, knockdown of other complex components would be predicted to have a similar effect on inhibiting reprogramming. RNAi of the Tip60/NuA4 complex homologs *trr-1*/TRRAP, *mys-1*/Esa1, and *ssl-1*/p400 does not appear to have any effect in preliminary experiments (data not shown).

SUMOylation may be required for gonad-to-intestine transorganogenesis

One particularly intriguing result from the RNAi screen of chromatin-related factors was the identification of *smo-1*, the sole *C. elegans* SUMO gene (Fig 7). SUMO is a small, ubiquitin-like modifier that can be covalently attached to proteins to modulate their activity or function in a highly reversible manner (Johnson, 2004). The SUMO pathway is well conserved and has numerous roles in development, acting in the nucleus to modulate the activity of transcription factors and chromatin modifiers, as well as acting on proteins in the cytoplasm (Broday, 2017; Lomelí and Vázquez, 2011). In *C. elegans,* SUMO is required for proper morphogenesis of the somatic gonad, germ line, and vulva, partially through sumoylation of LIN-11 (Broday et al., 2004), and has a conserved role in Hox gene regulation through control of Polycomb group activity (Zhang et al., 2004). SUMO has also

been found to play a role in pharynx development (Huber et al., 2013), cytoskeletal assembly (Kaminsky et al., 2009), and other processes (Broday, 2017).

Worms exposed to *smo-1(RNAi)* were observed to have significant gonad defects and sterility in a heat shock independent manner (Fig 11A). In scoring for somatic gonad reprogramming phenotypes, L4440 RNAi controls showed 100% (n = 50) transorganogenesis efficiency while *smo-1(RNAi)* had 14% (n = 50) with an ectopic lumen and an additional 54% showed some signs of reprogramming in the gonad (Fig 10). 34% had no detectable ectopic IFB-2::CFP in the gonad 24 hours after ectopic *elt-7* expression (Fig 11B). Ectopic gut expression in the pharynx was similar to controls.

One possible mechanism by which *smo-1(RNAi)* could interfere with reprogramming is by inhibiting the heat shock response and therefore preventing expression of the *hsp::elt-7* transgene. Induction of the heat shock response was measured by time lapse imaging of GFP fluorescence in the pharynx of a *hsp::gfp*-expressing strain that was treated with either control or *smo-1* RNAi and subjected to heat shock at the L4 stage. Quantification of GFP fluorescence indicates that on average there is no significant difference in the dynamics of heat shock response activation following knockdown of *smo-1* (Fig S1).

Another possible mechanism for the inhibition of somatic gonad-to-intestine transorganogenesis is through regulation of PHA-4, which is a factor necessary during embryogenesis for ELT-7 directed reprogramming of the pharynx (Riddle et al., 2016). Observation of a *pha-4::gfp* expressing strain treated with control or *smo-1* RNAi did not reveal any obvious differences in global *pha-4* expression, however wild type *pha-4::gfp* expression patterns in the gonad are dynamic and highly stage dependent, and it was unclear if subtle changes occurred within the developing gonad (data not shown).

Heat shock-independent activation and inhibition of transorganogenesis by *pyp-1* RNAi

A unique RNAi screening result was observed following knockdown of the inorganic pyrophosphatase gene *pyp-1* (Fig 7). The *pyp-1* gene encodes three protein isoforms that are orthologous to human PPA1 and *Drosophila* NURF-38 (Guindon et al., 2010; Ko et al., 2007; Ruan et al., 2008). *pyp-1* is known to be expressed in the intestine, nervous system, and coelomocytes, is required for larval development and intestinal function, and *pyp-1(RNAi)* can cause mislocalization of gonads and late larval arrest (Gaiser et al., 2009; Ko et al., 2007). As an ortholog of NURF-38, PYP-1 is also predicted to participate in nucleosome remodeling (Gdula et al., 1998).

Worms treated with *pyp-1(RNAi)* showed slow larval development and gonad defects, which is in agreement with previously reported phenotypes (Gaiser et al., 2009; Ko et al., 2007) (Fig 13A). In RNAi controls, 100% (n = 50) of *hsp::elt-7* expressing worms displayed the full ectopic lumen reprogramming phenotype as usual, while in *pyp-1(RNAi)* treated worms only 56% (n = 36) contained an ectopic lumen and 33% showed some ectopic gut formation (Fig 12). 11% of heat shocked *pyp-1(RNAi)* worms had no ectopic IFB-2::CFP in the somatic gonad (Fig 13A). Unexpectedly, *pyp-1(RNAi)* treated worms that were not subjected to heat shock also had ectopic *elt-2p::gfp* and *ifb-2::cfp* expression exclusively in the somatic gonad (Fig 12), with 35% (n = 20) displaying some ectopic gut expression and 40% displaying well-formed ectopic lumens, although more condensed in size and weaker CFP intensity than typically observed in controls (Fig 13B). In a control strain that does not contain any heat shock transgenes, *pyp-1(RNAi)* did not alter wild type gut marker expression patterns (Fig 12).

In preliminary experiments, a *hsp::gfp* strain treated with *pyp-1(RNAi)* at 20^0C showed GFP in the intestine, somatic gonad, coelomocytes, and excretory cell; GFP was not observed in the pharynx (data not shown). These results suggest that knockdown of inorganic pyrophosphatase expression can stochastically activate the heat shock response at 20^0C in select tissues, and therefore can induce heat shock-independent of *hsp::elt-7* expression and reprogramming of the somatic gonad. Additionally, *pyp-1* appears to play an important role in proper gonad development and competency for ELT-7 directed transorganogenesis.

Discussion

Redeployment of a robust gene regulatory network

Although the ELT-7 transcription factor is the most potent initiator of somatic gonad-to-intestine transorganogenesis, it does not appear to be absolutely required for this process. Being one of the smallest GATA factors involved in intestine differentiation, it was initially hypothesized that this property enables ELT-7 to reach deeper into compact chromatin to activate intestine genes. However, ectopic END-3 expression remains capable of reprogramming cells without an intact endogenous *elt-7* gene. ELT-7 therefore does not appear to possess any unique biomolecular function and is not solely responsible for facilitating cell fate conversions.

The intestine gene regulatory network (GRN) as a whole may be the key to driving cellular reprogramming in post-embryonic and fully differentiated cell types. During normal embryonic development, the endoderem GRN is deployed through a feed-forward transcription factor cascade and terminal differentiation is established through multiple

positive feedback loops that are thought to “lock-down” intestine cell fate (Maduro, 2015; Sommermann et al., 2010). Mapping of the spatial and temporal patterns of ectopic *elt-2p::gfp* and *ifb-2::cfp* expression following *hsp::elt-7* induction suggests that the sequence of events in somatic gonad-to-intestine transorganogenesis is related to normal embryonic developmental events. Much of the intestine GRN may therefore be redeployed within the context of an unrelated tissue during reprogramming into intestine, even when one of network nodes has been deleted.

The limits of the intestine GRN to redirect cell fate have not yet been reached and will require further experiments testing the loss of endoderm factors on the ability to reprogram tissues. Of particular interest is whether END-3 would be capable of initiating somatic gonad-to-intestine transorganogenesis without either of the downstream genes *elt-7* and *elt-2*. Unfortunately, initial experiments testing reprogramming in *hsp::end-3; elt-7(-); elt-2(RNAi)* animals were inconclusive owing to developmental defects that precluded analysis of somatic gonad reprogramming effects. A more sophisticated genetic approach could be utilized to generate a conditional or tissue-specific *elt-7(-); elt-2(-)* double knock out to test for late larval reprogramming ability.

Nuclear factors that directly affect the ability of ELT-7 to reprogram tissues

In this chapter I presented the results of screening a total of 338 genes by RNA interference from two closely related sets of candidate factors chosen based on their functions as direct DNA binding transcription factors and chromatin-related factors. The majority of results were negative and less than 10% were considered hits for modulating reprogramming efficiency. Initial retesting indicated that most hits did not replicate well

(data not shown), and ultimately only ~1% of genes tested gave reproducible results that were followed up on experimentally. Although genes upregulated after ectopic ELT-7 are strongly enriched for GATA and Forkhead transcription factor binding sites, knockdown of any one single factor from either family does not inhibit transorganogenesis.

A clear difference in the ability of ELT-7 to reprogram tissues was observed in RNAi of *epc-1*/Enhancer of Polycomb 1. Histone acetyltransferases have diverse functions but are best known for their roles is transcriptional activation and it is plausible that the *C. elegans* Tip60/NuA4-like complex could be acting in the somatic gonad to establish a chromatin state that is more permissible to reprogramming into intestine (Carrozza et al., 2003; Doyon et al., 2004). Disruption of this complex therefore would result in a reduced ability to undergo somatic gonad-to-intestine transorganogenesis. However, *epc-1(RNAi)* also causes larval arrest and embryonic lethality, as well as potentially interacting with the heat shock response pathway to stochastically activate *hsp::elt-7* and *elt-2p::gfp* expression. These complex phenotypic interactions need to be further deconvolved to determine whether *epc-1* is directly functioning to promote reprogramming. Additional experiments focusing on plasticity of early embryos or on transorganogenesis in a heat shock independent system may clarify the contribution of *epc-1* to ELT-7-directed cellular reprogramming.

The SUMO pathway is well conserved from yeast to humans and can affect transcription in a multitude of ways, such as direct modification of transcription factors that may alter DNA binding activity (Ward et al., 2013), or sumoylation of histone deacetylases, which reduces transcriptional repression activity (Zhang et al., 2004). I found that inhibition of *smo-1*/SUMO reduces the susceptibility of the somatic gonad to forced reprogramming into intestine, suggesting that sumoylation plays a role in promoting transorganogenesis.

Conversely, suppression of SUMO2 can promote the generation of human induced pluripotent stem cells, indicating that sumoylation can also serve as a roadblock to reprogramming (Borkent et al., 2016). The nature of the relationship between SUMO and transorganogenesis in *C. elegans*, however, remains ambiguous, as *smo-1(RNAi)* also results in gonad development defects and inhibition may disrupt multiple cellular processes. The specificity of this relationship could be further strengthened through targeting other factors in the SUMO pathway, such as the SUMO conjugating enzyme UBC-9. Additionally, a recent study identified 874 SUMO targets in the worm (Drabikowski et al., 2018), which may contain more specific and promising candidates for regulation of transorganogenesis.

The presence of ectopic intestine-like lumens in non-heat shocked *pyp-1(RNAi)* treated worms was wholly unexpected and unique. This appears to be the result of spontaneous activation of the heat shock response due to loss of pyrophosphatase activity, a previously unreported connection in *C. elegans*. The molecular mechanism that relates PYP-1 function to the heat shock response is unknown. Furthermore, the role of PYP-1 in gonad development is not well characterized and it is unclear why *pyp-1(RNAi)* activates the heat shock response in a tissue-specific pattern. The interplay between the effects of *pyp-1(RNAi)* on cellular reprogramming, induction of the heat shock response, development of the gonad, and potential nucleosome remodeling is confounding. Further research into the molecular functions of PYP-1 in *C. elegans* is needed and may elucidate the mechanism by which it modulates transorganogenesis of the somatic gonad into intestine.

Limitations of the RNAi screen

RNA interference is a powerful experimental tool in *C. elegans* for rapidly testing large numbers of genes for a particular phenotype (Ceron et al., 2007; Kamath et al., 2003; Rual et al., 2004). Ideally, a genome wide reverse genetic screen could be performed to identify factors that promote or prohibit reprogramming into intestine. However, a screen of such scale requires an essentially qualitative phenotypic output that can be rapidly scored and/or automated. The goals of the RNAi screens performed required manual scoring of often subtle phenotypes and therefore testing capacity was limited to a few hundred factors. Alternative genetic screening approaches have been devised and are in progress in the Rothman lab, including a loss-of-function mutagenesis screen to select for mutants that are resistant to ELT-7-induced developmental arrest, as well as a gain-of-function transposition-based screen to identify mutants in which ELT-7 can reprogram additional tissue types.

The RNAi screens described here were designed to test for factors that are required at the time of reprogramming. Worms were fed on RNAi expressing bacteria from the L1 to early L4 stage under the assumption that approximately 2 days of feeding is sufficient for strong systemic knockdown of gene expression. Additionally, RNAi treatment was started at the L1 stage in order to minimize embryonic lineage effects that might prevent the specification or development of the somatic gonad, resulting in gross morphological defects that could confound interpretation of effects on reprogramming or yield false positive hits. However, some proteins may be produced during embryogenesis that function during reprogramming, in which case RNAi starting at L1 may be insufficient to inhibit their activity, thereby resulting in what are essentially false negatives. In order to overcome this limitation, it would be necessary to perform maternal RNAi knockdowns, starting at the L4 stage in the P0

generation and then perform testing for reprogramming efficiency in the F1 generation. Such a screening approach would be moderately more complex but would more thoroughly test whether a certain factor is involved in reprogramming and may yield a higher positive hit rate.

Conclusions

The ability of the endoderm GRN to establish intestine cell fate is very robust, even in an ectopic cellular context. Despite significant disruption to the GRN itself or to a broad assortment of nuclear factors, ELT-7 remains capable of initiating transdifferentiation of the pharynx and transorganogenesis of the somatic gonad into intestine-like tissue with near-total efficiency. The most meaningful inhibitory effects on reprogramming are seen with chromatin-related factors. These results are consistent with the hypothesis that the expression of some factor in the somatic gonad acts to establish a cellular context that is permissive to reprogramming to an intestine fate. The biomolecular nature of this priming remains elusive, but it may involve the establishment of a chromatin state that can be readily accessed by intestine factors.

Materials and Methods

Worm strains and maintenance

C. elegans strains were grown at 20^0C on NGM media seeded with OP50 bacteria as previously described (Brenner, 1974; Stiernagle, 2006) JR3666 rrIs1[elt-2p::gfp]; kcIs6[ifb-2::cfp], JR3642 wIs125[hsp::elt-7 + pRF4 (rol-6(su1004))]; rrIs1[elt-2p::gfp]; kcIs6[ifb-2::cfp], JR3600 wIs76[hs::end-3 + Rol]; rrIs1[elt-2p::gfp]; kcIs6[ifb-2::cfp], JR3785 elt-7(ok835) V; wIs76[hs::end-3 + Rol]; rrIs1[elt-2p::gfp]; kcIs6[ifb-2::cfp], CL2070 dvIs70[hsp-16.2p::gfp + rol-6(su1006)].

Construction of *hsp::end-3; elt-7(ok835)* strain

The JR3600 strain was crossed with RB948. Worms possessing the ok835 deletion were tracked using PCR amplicon length genotyping. Forward primer: TGC TCA CAC TGC TCA AC; reverse primer: CTT GAC GCC GCT TTC GAG TA, which amplify an 805 bp region of *elt-7(+)* and a 307 bp region of *elt-7(ok835)*. PCR protocol: 95^0, 3:00; [95^0, 0:30; 54^0, 0:30; 70^0, 1:00] x35; 70^0, 5:00.

Enrichment of predicted transcription factor binding sites

Identification of transcription factors with enrichment of binding sites for genes upregulated following ectopic *elt-7* expression was done using the web-based tool oPOSSUM version 3.0 (Kwon et al., 2012; Sui et al., 2005; Sui et al., 2007). For input, the top 2,000 transcripts upregulated 3 hours post heat shock were used, as detected from RNA-seq analysis [see Chapter 3 for more detail]. All 26,473 genes in the oPOSSUM database were used as the background list for each analysis. Multiple analyses were run using varying parameters for the length of upstream and downstream sequences, matrix score threshold,

and conservation cutoff. To compare the enrichments between analyses, z-scores were normalized to the most significant value.

RNAi screening

Bacterial RNAi cultures were obtained from frozen glycerol stocks originating from either the Rothman lab collection, the Vidal RNAi library (Rual et al., 2004), or the Ahringer RNAi library (Kamath et al., 2003), with preference for the more former of the sources if multiple stocks were available, and streaked onto large LB plates containing ampicillin (Amp) and tetracycline (Tet). Single bacterial colonies were selected to inoculate 2 mL of LB spiked with 50 µg/mL Amp and 5µg/mL Tet and incubated at 37^0C overnight in a shaker. Overnight culture was added to fresh LB with Amp at a 1:10 ratio and incubated in a 37^0C shaker for 4 hours. RNAi cultures were spiked with 1 mM IPTG and then ~100 µL of bacteria were seeded onto small RNAi NGM plates. Plates were allowed to dry and then incubated overnight at 37^0C or for 2 days at room temperature.

Worms were prepared starting by transferring a ~1 cm^2 chunk from a starved plate of JR3642 strain worms to a fresh, medium NGM plate with an OP50 bacterial lawn. After 3 days, or when egg laying was maximal, the plate was used for an egg prep. Briefly, worms and eggs were rinsed with M9 or gently scraped off of the plate with a bent glass capillary tube into a 1.5 mL tube, pelleted, and rinsed 3x, gently resuspending worms between washes. Worms were then resuspended in freshly prepared egg isolation solution (1 mL household bleach, 0.25 mL 10N NaOH, 3.75 mL MiliQ water) and rocked until most worms were lysed. Remaining embryos were pelleted and washed at least 3x, and left to hatch in M9 on a rotator at 15^0C overnight. Approximately 100 synchronized, starved L1 worms were seeded onto each RNAi plate and incubated at 20^0C. After 42-45 hours of feeding on RNAi, when worms

had reached early L4 stage, worms were heat shocked by placing plates in a 33^0C incubator for 30 minutes and then returned to 20^0C. Worms were scored 24 hours post heat shock.

Imaging and scoring

Worms were transferred to microscope slides with a 4% agar pad, immobilized with ~1 µM levamisole and then viewed on a Nikon Eclipse Ti inverted fluorescence microscope; images were captured with a Hamamatsu flash Orca 2.8 camera. Scoring was based on the presence and morphology of IFB-2::CFP in the somatic gonad and the major phenotypic categories were as follows: worms with a transorganogenesis phenotype typical of controls, where the ectopic IFB-2::CFP forms a secondary intestine-like lumen structure were scored "lumen"; if some ectopic IFB-2::CFP was observed in the somatic gonad but did not form a continuous lumen structure, worms were scored "some ect."; worms with an ectopic lumen that was wider and had a more diffuse IFB-2::CFP signal, typically associated with reprogramming in late L4s, were scored "wide lumen"; if IFB-2::CFP was not seen in the uterus but was clearly present in the spermatheca, commonly associated with reprogramming in young adult hermaphrodites, worms were scored "spth"; and if no ectopic IFB-2::CFP was observed in the somatic gonad, worms were scored "no ect." Some additional categories were included for unusual phenotypes, such as worms that had ruptured ("exploded"), or worms with notable IFB-2::CFP in other regions or tissues (ie: some CFP in the head, "phx,gon,head"; or CFP in the excretory cell, "excretory cell"). In nearly all conditions IFB-2::CFP was present in the pharynx. elt-2p::GFP expression was always monitored during scoring but was not used as the basis for assigning phenotypic scores except for in rare cases ("lumen,super GFP").

Fig 1.

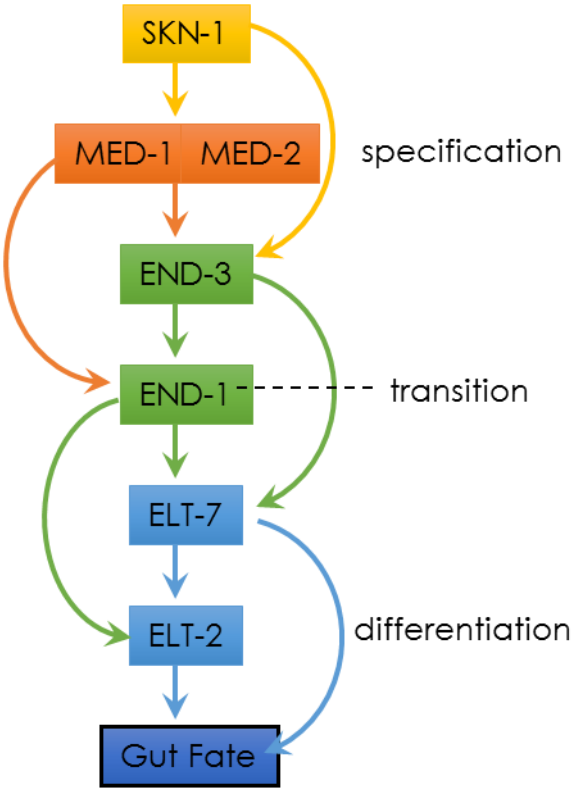

SKN-1
MED-1
MED-2
specification
END-3
END-1
transition
ELT-7
ELT-2
differentiation
Gut Fate

Fig 1. The recursive feed-forward regulatory logic for intestine development in *C. elegans*.

The maternally derived SKN-1 protein activates the *med-1* and *med-2* genes to specify the EMS cell. These genes then activate *end-3* and *end-1* in the E cells. These transcription factors in turn activate the terminal differentiation transcription factors *elt-7* and *elt-2* which directly promote the transcription of thousands of intestine-specific genes. This transcription factor cascade is especially robust owing to the ability of each factor to activate the next two factors in the sequence, which allows for successful intestine development in spite of transcriptional perturbations or even complete loss of one of the factors. (Adapted from Sommermann et al., 2010).

Fig 2.

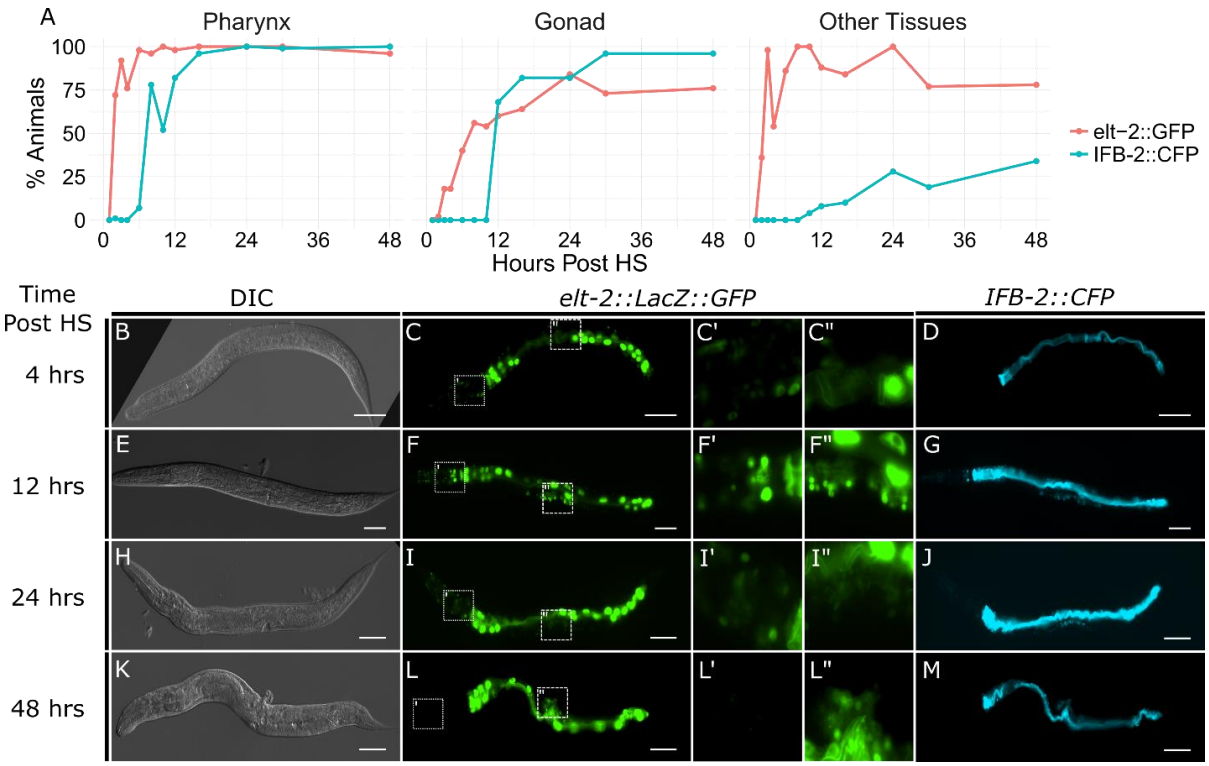

A
Pharynx
Gonad
Other Tissues
% Animals
100
75
50
25
0
0
12
24
36
48
Hours Post HS
elt−2::GFP
IFB-2::CFP
Time Post HS
DIC
elt-2::LacZ::GFP
IFB-2::CFP
4 hrs
12 hrs
24 hrs
48 hrs
B
C
C'
C''
D
E
F
F'
F''
G
H
I
I'
I''
J
K
L
L'
L''
M

Fig 2. Spatio-temporal patterns of intestine markers after ectopic ELT-7 expression in early L4 stage worms.

(A) Proportion of worms that show ectopic expression of the intestine-specific transgenic markers *elt-2::gfp* and *ifb-2::cfp*. The spatial expression for each marker was assessed as first detectable in the pharynx, gonad, and other tissues excluding endogenous intestine expression. (B-M) Representative fluorescence and DIC images of worms at progressive intervals following widespread activation of the *hsp::elt-7* transgene. Ectopic GFP expression in the pharynx is visible in most worms by 4 hours post heat shock (HS) (C'), and in the gonad of a majority of worms by 12 hours post HS (F"). The intensity of ectopic CFP in other tissues is considerably weaker and sporadic compared to pharynx and gonad expression. Scale bars, 50 μm. [A similar version of this figure was published in Riddle et al. 2016].

Fig 3.

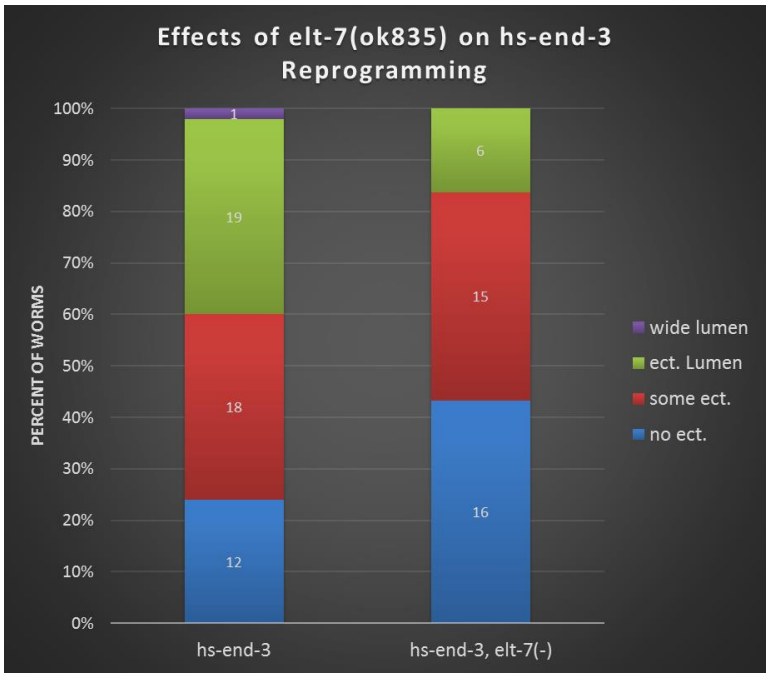

Effects of elt-7(ok835) on hs-end-3 Reprogramming
PERCENT OF WORMS
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1
19
18
12
6
15
16
hs-end-3
hs-end-3, elt-7(-)
wide lumen
ect. Lumen
some ect.
no ect.

Fig 3. Effects of *elt-7(ok835)* on *hsp::end-3* reprogramming.

The *elt-7(ok835)* deletion mutation was crossed into the *hsp::end-3* strain, which is capable of inducing somatic gonad-to-intestine transorganogenesis. Loss of *elt-7* causes a slight decrease in reprogramming efficiency but this difference is not statistically significant. This result indicates that although ectopic expression of *elt-7* is sufficient to induce highly efficient transorganogenesis, the ELT-7 transcription factor is not essential for this reprogramming process. See Materials and Methods for details on scoring.

Fig 4.

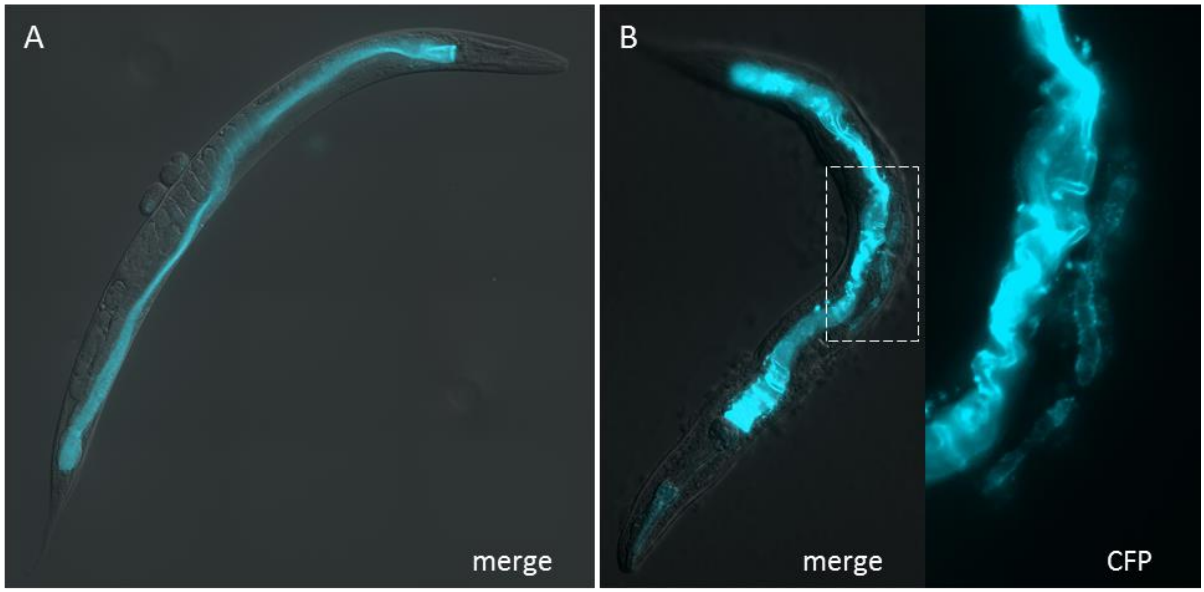


Fig 4. ELT-7 directed reprogramming of the pharynx and transorganogenesis of the somatic gonad into intestine-like tissue.

(A) Control adult worm with wild type expression of IFB-2::CFP in the intestine. (B) *hsp::elt-7* expressing worm 24 hours after heat shock with ectopic IFB-2::CFP in the pharynx and somatic gonad. Inset: reprogrammed somatic gonad, showing distinct ectopic intestine-like lumen formation.

Fig 5.

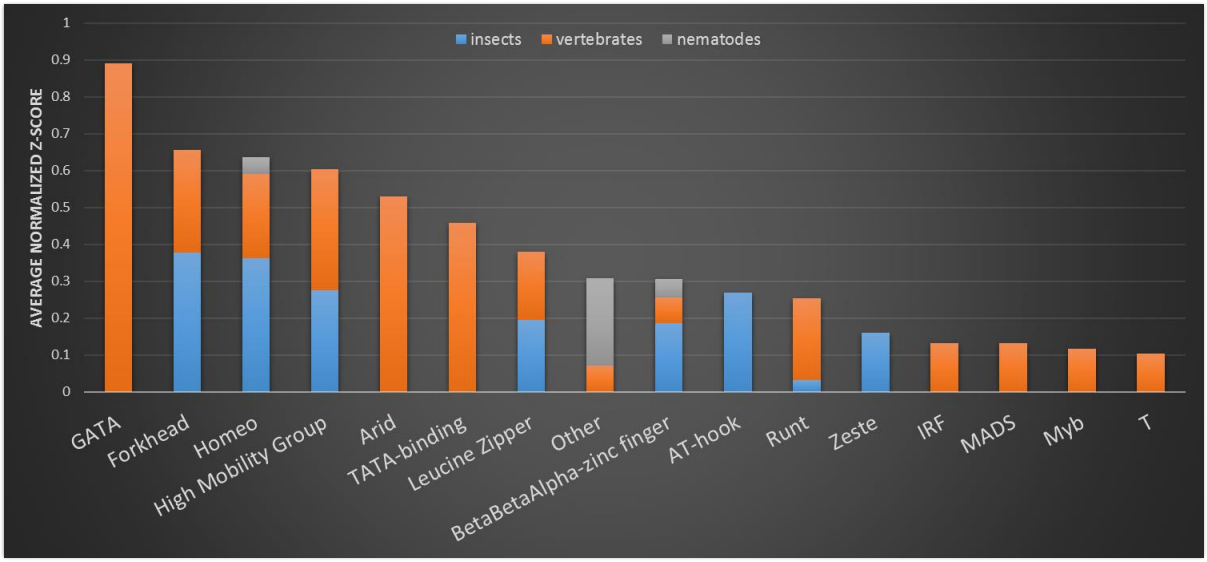

insects
vertebrates
nematodes
AVERAGE NORMALIZED Z-SCORE
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
GATA
Forkhead
Homeo
High Mobility Group
Arid
TATA-binding
Leucine Zipper
Other
BetaBetaAlpha-zinc finger
AT-hook
Runt
Zeste
IRF
MADS
Myb
T

Fig 5. Transcription factor families with enriched binding sites among genes upregulated 3 hours after ectopic ELT-7.

The oPOSSUM analysis tool was used to predict transcription factors that may be activated by ectopic *elt-7* expression, based on the enrichment of transcription factor binding sites (TFBS) in regulatory regions of genes that are upregulated at an early stage in transorganogenesis. Multiple oPOSSUM analyses were performed and TFBS enrichment scores were averaged to determine the transcription factor families with the most prevalent binding sites. Each bar also shows the phyla with TFBS predictions in the oPOSSUM database for transcription factor families. GATA factor binding sites are the most significantly enriched, possibly because of ELT-7 and ELT-2 transcriptional activity. *C. elegans* homologs from the top transcription factor families were prioritized for RNAi screening.

Fig 6.

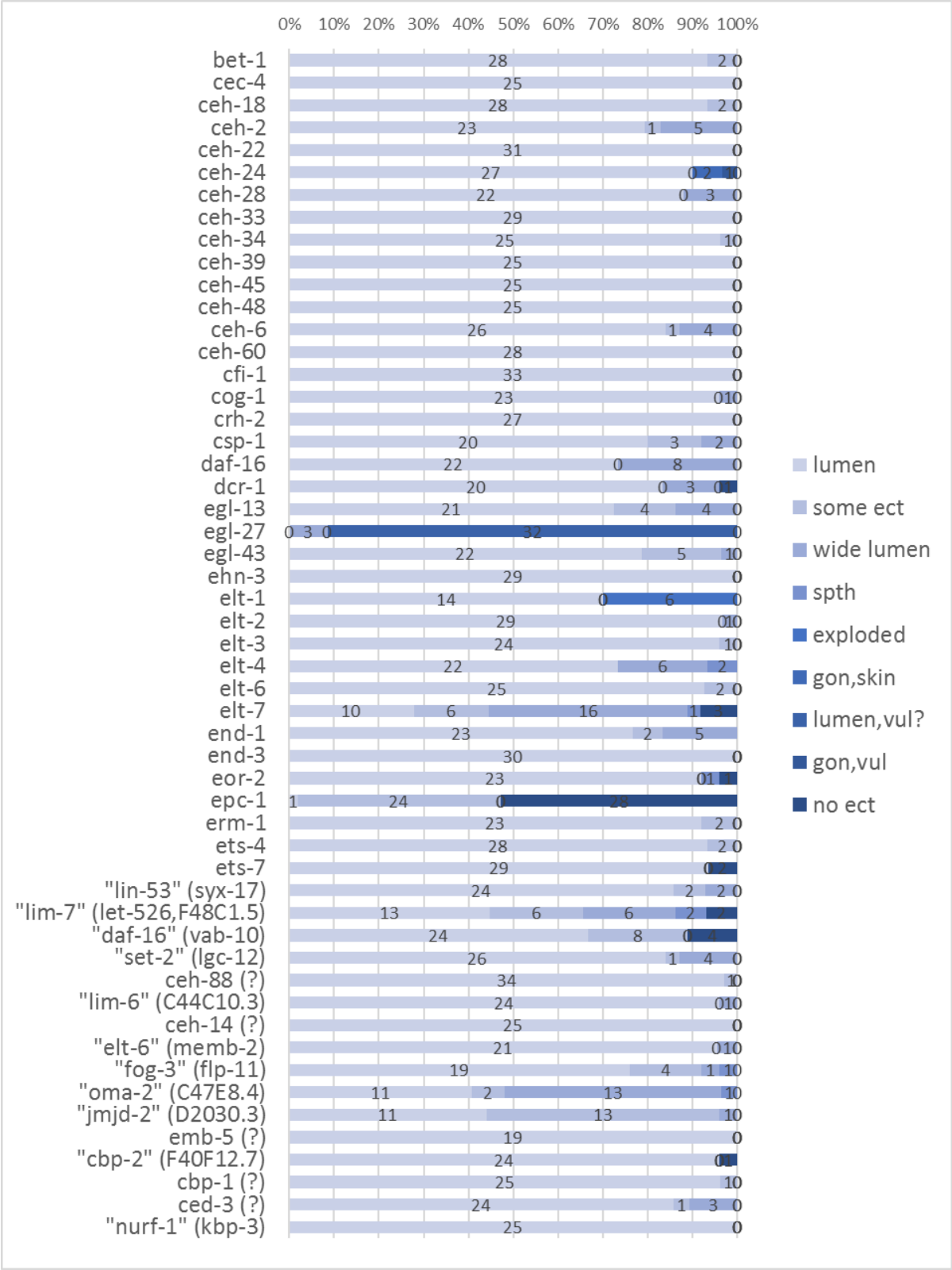

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
bet-1
cec-4
ceh-18
ceh-2
ceh-22
ceh-24
ceh-28
ceh-33
ceh-34
ceh-39
ceh-45
ceh-48
ceh-6
ceh-60
cfi-1
cog-1
crh-2
csp-1
daf-16
dcr-1
egl-13
egl-27
egl-43
ehn-3
elt-1
elt-2
elt-3
elt-4
elt-6
elt-7
end-1
end-3
eor-2
epc-1
erm-1
ets-4
ets-7
"lin-53" (syx-17)
"lim-7" (let-526,F48C1.5)
"daf-16" (vab-10)
"set-2" (lgc-12)
ceh-88 (?)
"lim-6" (C44C10.3)
ceh-14 (?)
"elt-6" (memb-2)
"fog-3" (flp-11)
"oma-2" (C47E8.4)
"jmjd-2" (D2030.3)
emb-5 (?)
"cbp-2" (F40F12.7)
cbp-1 (?)
ced-3 (?)
"nurf-1" (kbp-3)
lumen
some ect
wide lumen
spth
exploded
gon,skin
lumen,vul?
gon,vul
no ect

Fig 6. (continued)

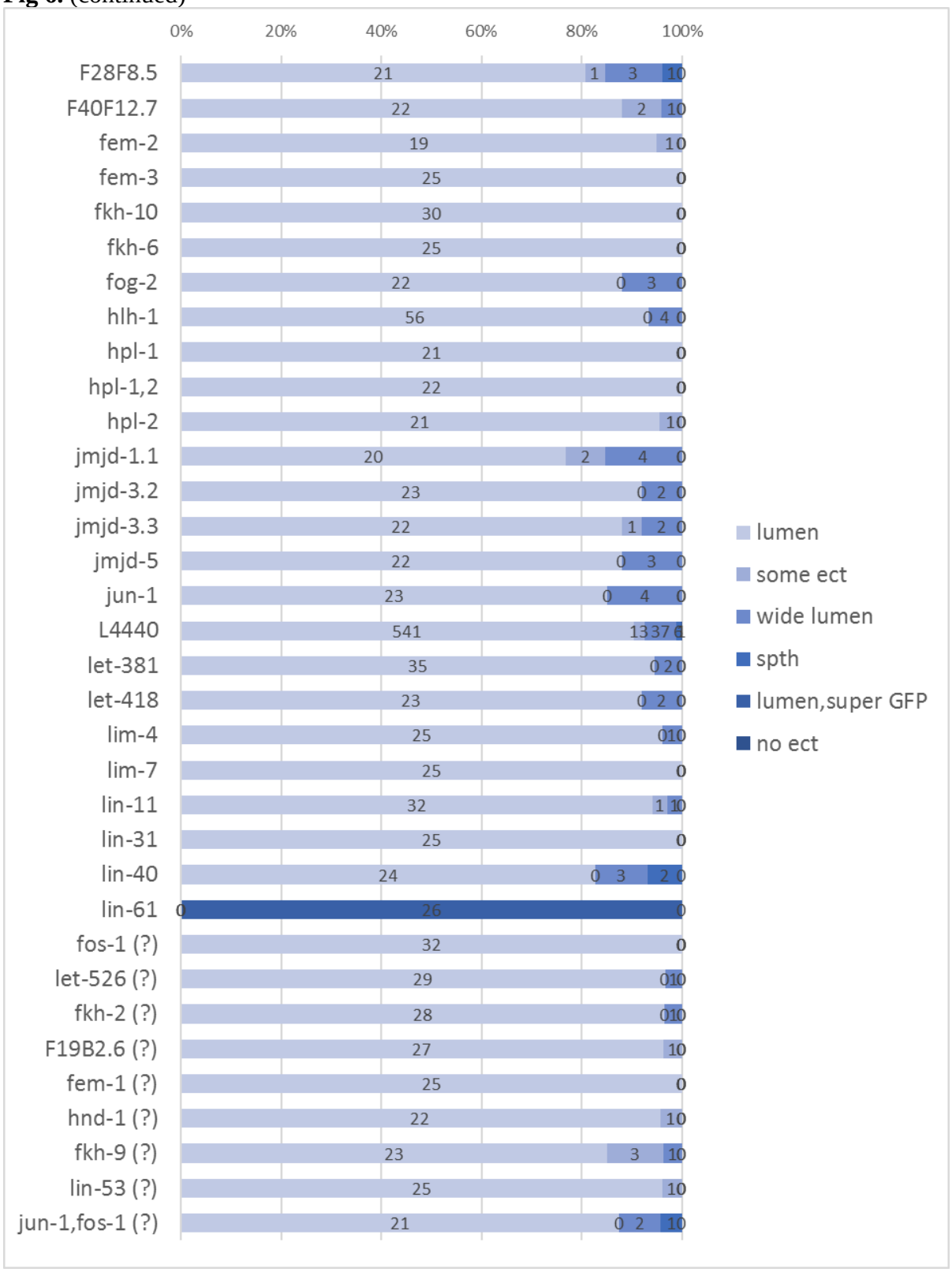


Fig 6. (continued)

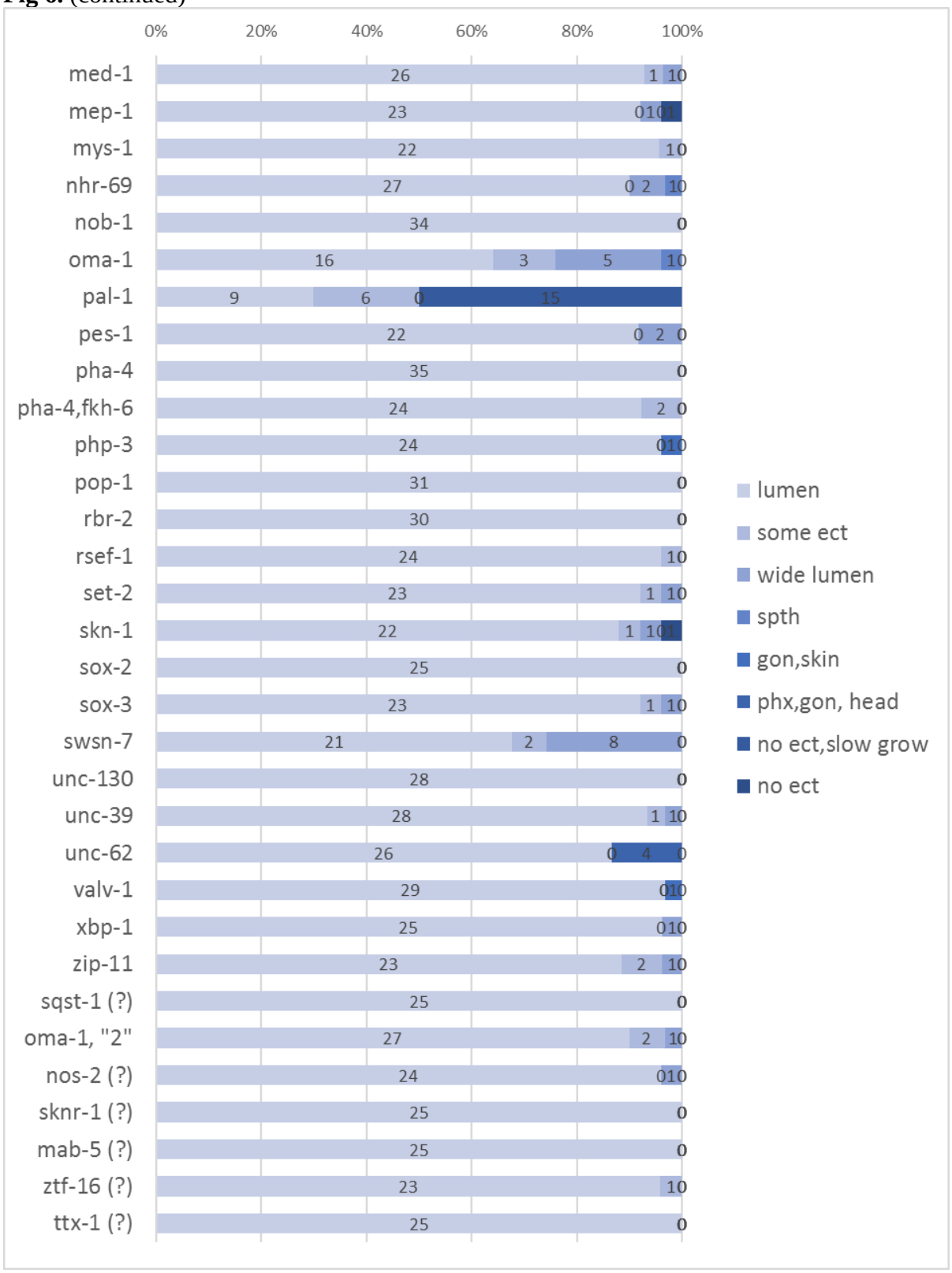


Fig 6. Transcription factor-based RNAi screening results.

Stacked bar charts showing the percentage of worms displaying each phenotype 24 hours post heat shock with n's centered for each category. Phenotypes are arranged in order from least to greatest effect on reprogramming, with darker colors indicating rare or strong effects. See L4440 for pooled control results. Genes for which the RNAi sequence was unable to be confirmed are marked with (?), and genes for which the RNAi sequence was determined to be incorrect are listed in "quotes" with the actual corresponding gene in (parentheses). See Materials and Methods section for details on scoring phenotypes.

Fig 7.

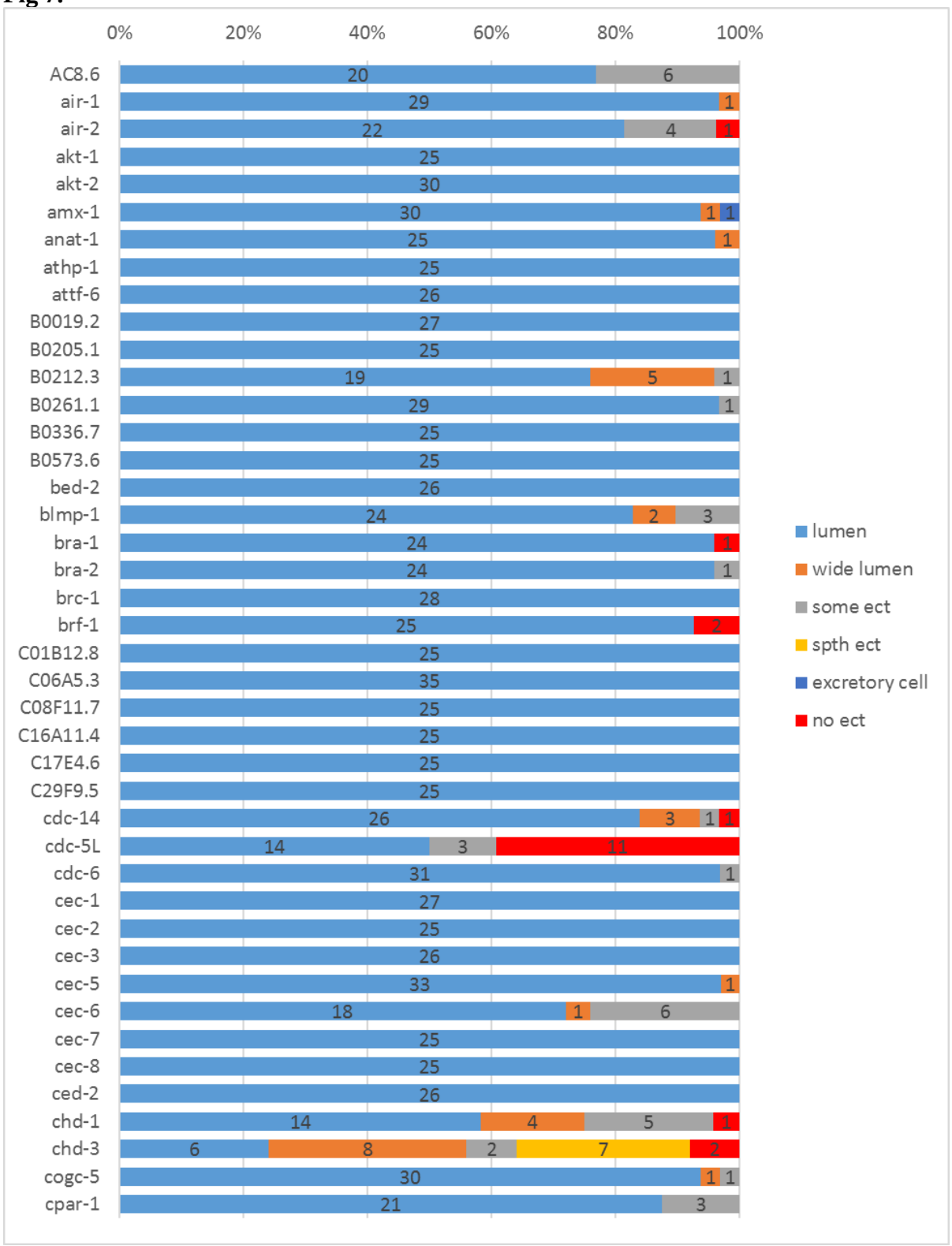

0%
20%
40%
60%
80%
100%
AC8.6 20 6
air-1 29 1
air-2 22 4 1
akt-1 25
akt-2 30
amx-1 30 1 1
anat-1 25 1
athp-1 25
attf-6 26
B0019.2 27
B0205.1 25
B0212.3 19 5 1
B0261.1 29 1
B0336.7 25
B0573.6 25
bed-2 26
blmp-1 24 2 3
bra-1 24 1
bra-2 24 1
brc-1 28
brf-1 25 2
C01B12.8 25
C06A5.3 35
C08F11.7 25
C16A11.4 25
C17E4.6 25
C29F9.5 25
cdc-14 26 3 1 1
cdc-5L 14 3 11
cdc-6 31 1
cec-1 27
cec-2 25
cec-3 26
cec-5 33 1
cec-6 18 1 6
cec-7 25
cec-8 25
ced-2 26
chd-1 14 4 5 1
chd-3 6 8 2 7 2
cogc-5 30 1 1
cpar-1 21 3
lumen
wide lumen
some ect
spth ect
excretory cell
no ect

Fig 7. (continued)

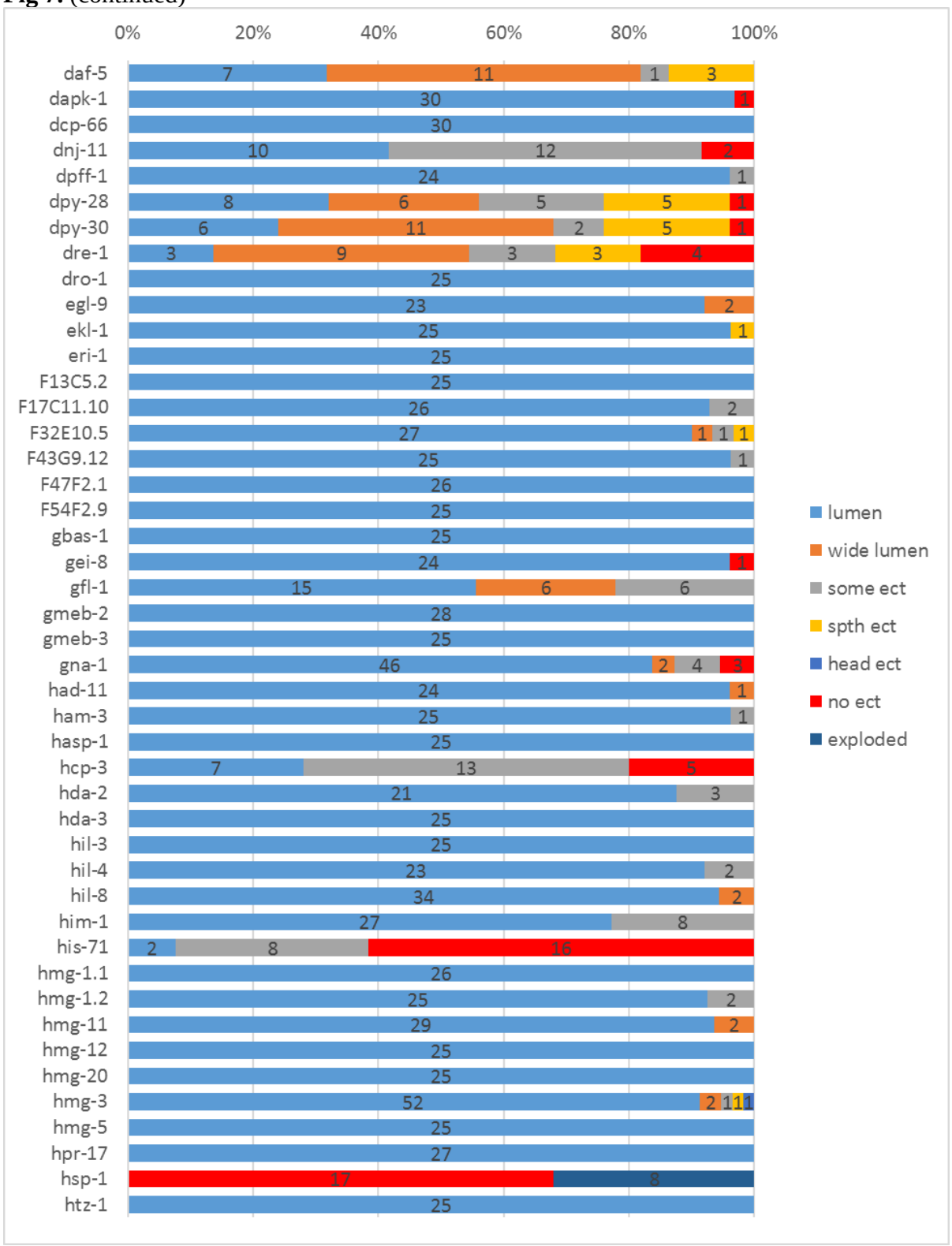


Fig 7. (continued)

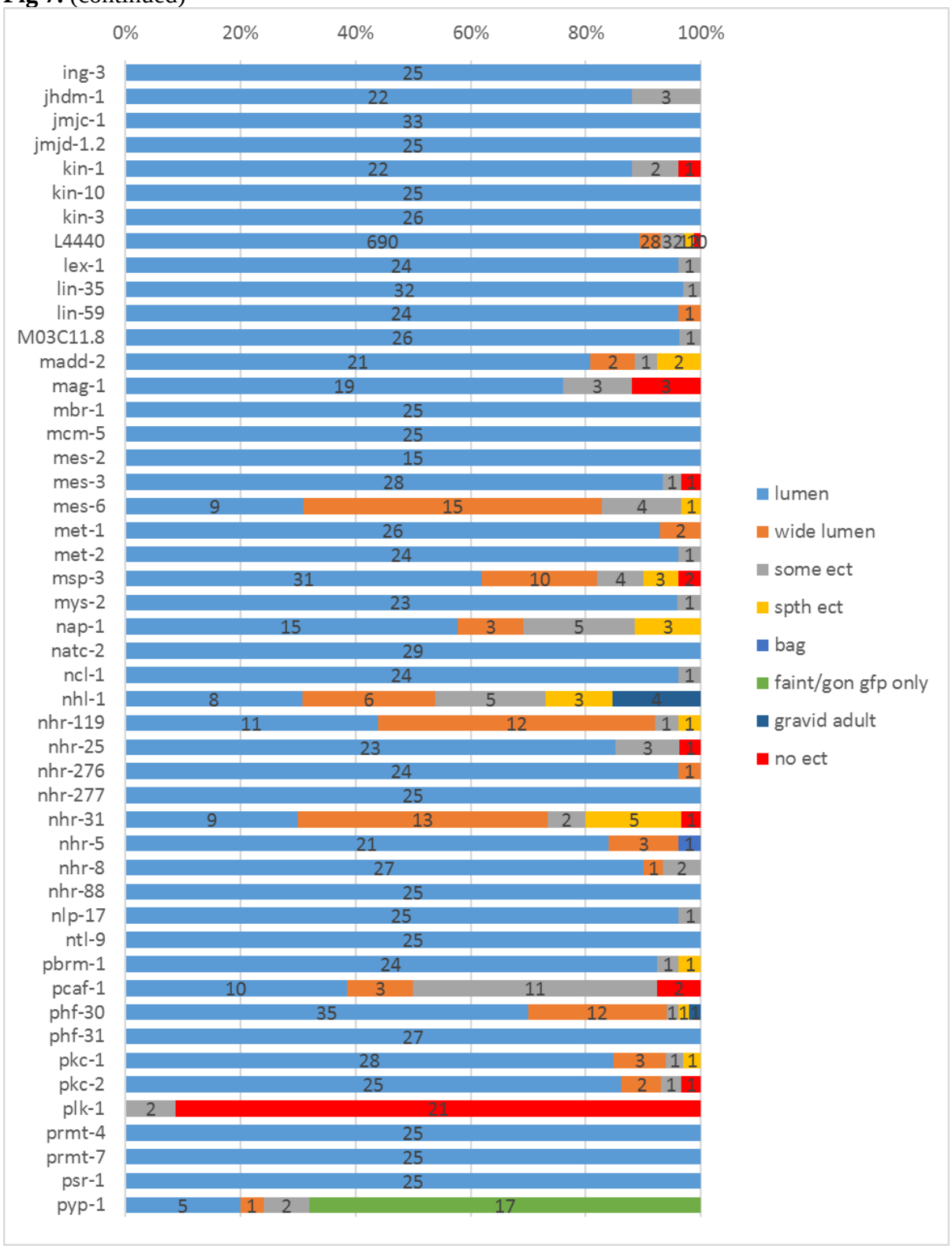


Fig 7. (continued)

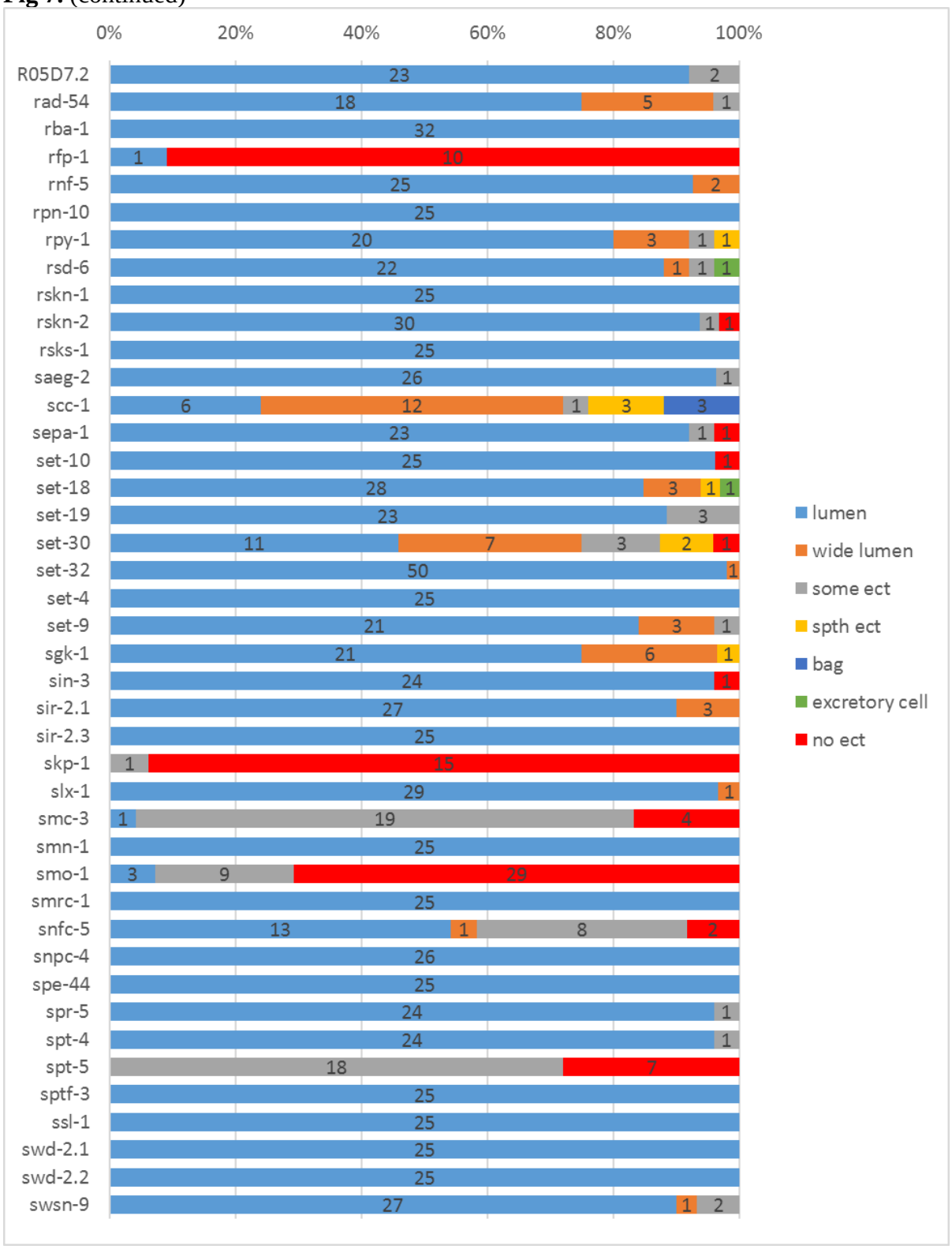


Fig 7. (continued)

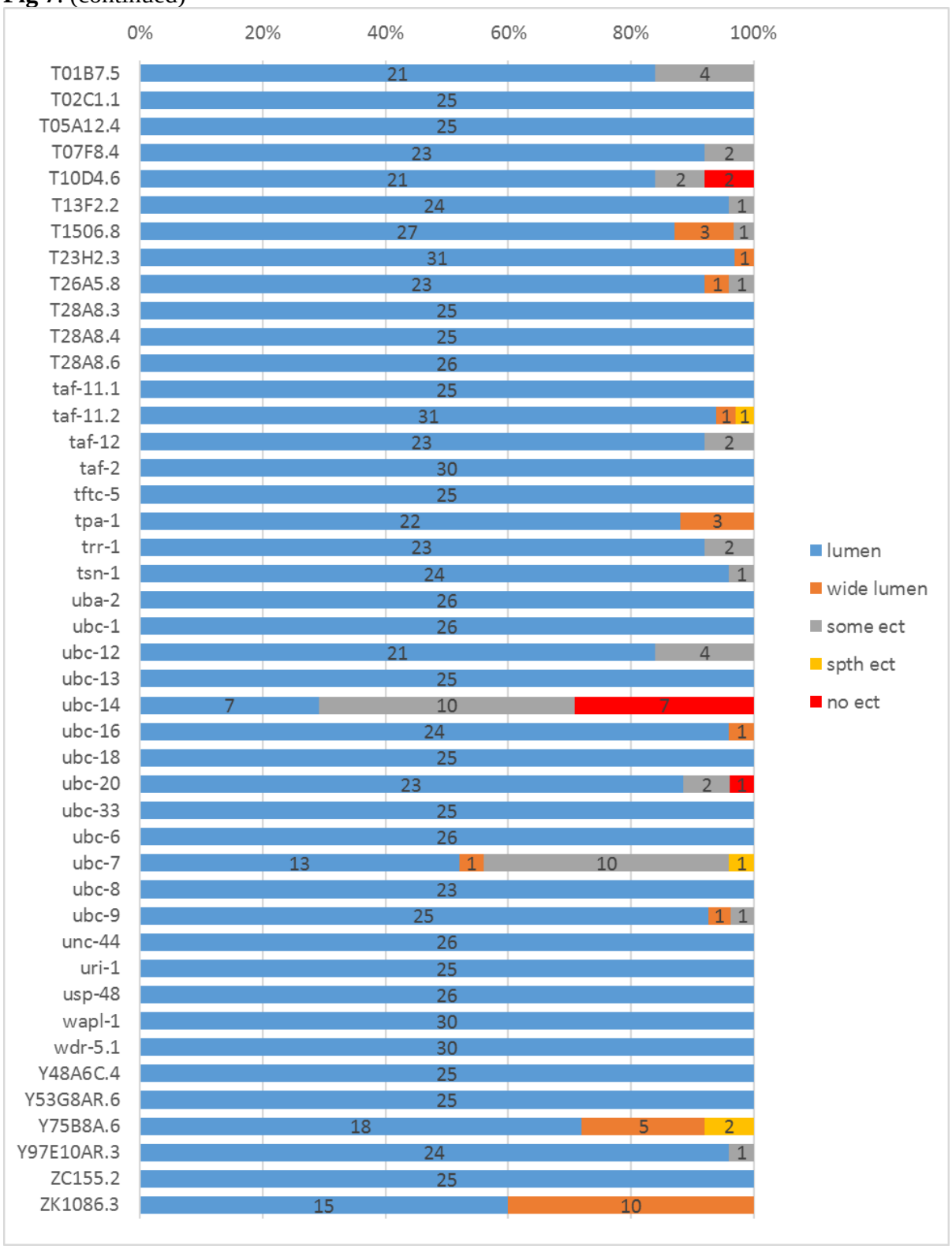


Fig 7. Chromatin factor-based RNAi screening results.

Stacked bar charts showing the percentage of worms displaying each phenotype 24 hours post heat shock with n's centered for each category. Phenotypes are arranged in order from least to greatest effect on reprogramming. See L4440 for pooled control RNAi results. The most significant hits for factors that contribute to transorganogenesis are RNAi's with the lowest percentage of worms with the "lumen" phenotype (blue) and/or the greatest percentage of worms with the "no ect" phenotype (red). Note that "wide lumen" (orange) and "spth" (yellow) phenotypes are commonly associated with experimental variability and do not necessarily indicate significant alterations in reprogramming efficiency (see Methods for scoring details).

Fig 8.

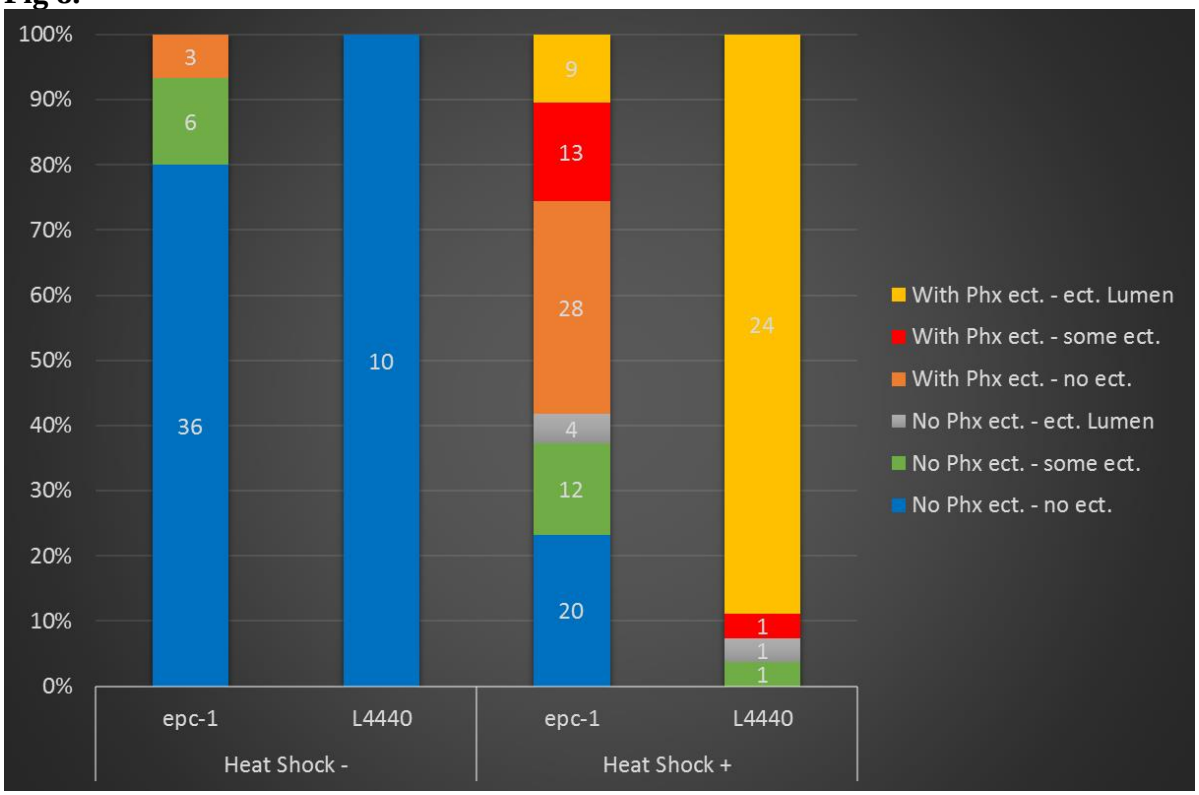

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
3
6
36
10
9
13
28
4
12
20
24
1
1
1
epc-1
L4440
epc-1
L4440
Heat Shock -
Heat Shock +
With Phx ect. - ect. Lumen
With Phx ect. - some ect.
With Phx ect. - no ect.
No Phx ect. - ect. Lumen
No Phx ect. - some ect.
No Phx ect. - no ect.

Fig 8. Knockdown of *epc-1* differentially inhibits pharynx and somatic gonad reprogramming and causes heat shock independent intestine marker expression.

The frequencies of worms exhibiting somatic gonad reprogramming phenotypes with and without accompanying reprogramming within the pharynx are shown for *epc-1(RNAi)* treated and control RNAi worms 24 hours after ectopic ELT-7. In the strain carrying the *hsp::elt-7* transgene, some ectopic intestine marker gene expression was observed in *epc-1(RNAi)* treated animals that were not heat shocked.

Fig 9.

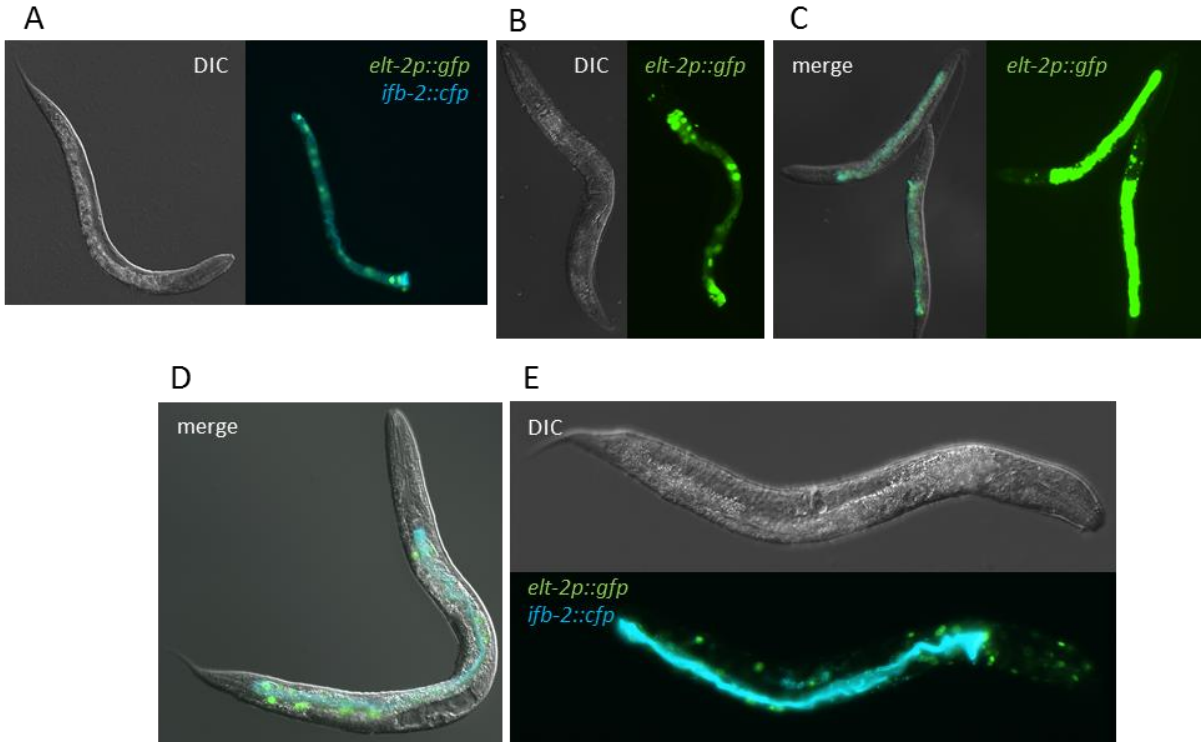

A
DIC
elt-2p::gfp
ifb-2::cfp
B
DIC
elt-2p::gfp
C
merge
elt-2p::gfp
D
merge
E
DIC
elt-2p::gfp
ifb-2::cfp

Fig 9. *epc-1(RNAi)* has complex phenotypic effects on *elt-7* directed reprogramming. (A) Representative worm after feeding on *epc-1(RNAi)* for two days starting at L1, displaying slow/arrested larval development. (B) Worm after feeding on *epc-1(RNAi)* for three days with ectopic *elt-2p::gfp* expression at continuous 20^0C incubation. (C) 24 hours post heat shock of a worm as in panel A displaying weak ectopic *elt-2p::gfp* expression and no detectable ectopic *ifb-2::cfp* expression. (D) Worm after feeding on *epc-1(RNAi)* for two days starting at ~L2, arresting later in development with a poorly formed gonad. (E) 24 hours post heat shock of a worm as in panel (D) with more significant ectopic intestine expression compared to (C). Brightness and contrast have been adjusted to visualize weak ectopic fluorescence patterns.

Fig 10.

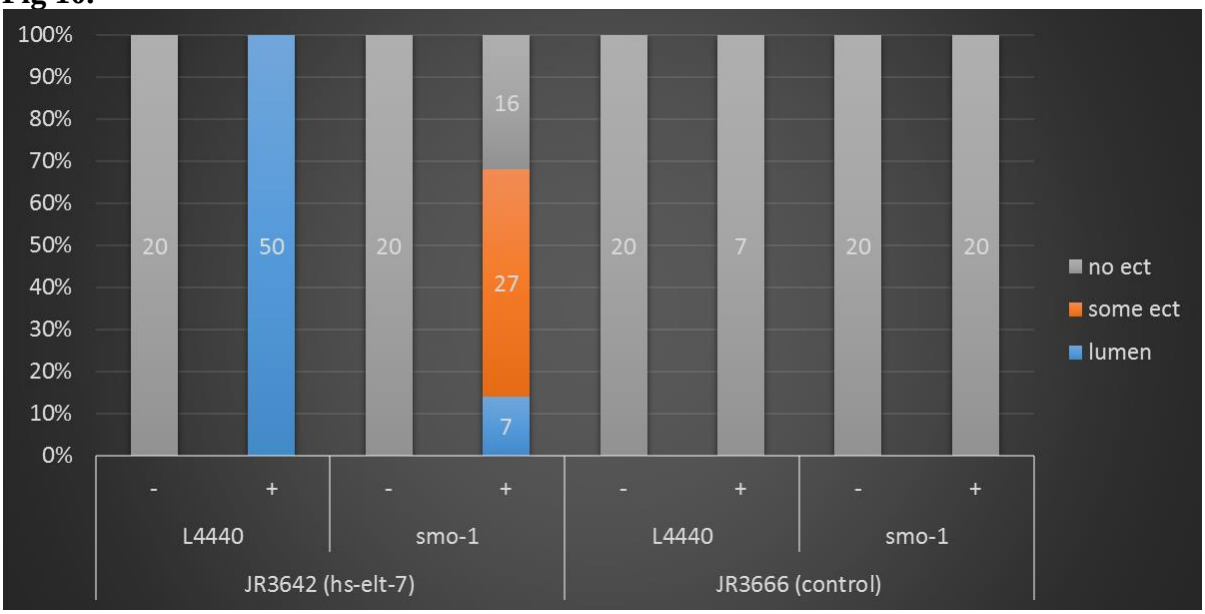

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
20
50
20
16
27
7
20
7
20
20
no ect
some ect
lumen
-
+
-
+
-
+
-
+
L4440
smo-1
L4440
smo-1
JR3642 (hs-elt-7)
JR3666 (control)

Fig 10. *smo-1(RNAi)* dramatically inhibits ELT-7 directed transorganogenesis. Knockdown of *smo-1*/SUMO results in a decrease in reprogramming efficiency of the somatic gonad driven by ectopic ELT-7. SUMO is a small ubiquitin-like modifier that can alter the activity of numerous factors and these results suggest that the SUMO pathway is important for effective reprogramming of the uterus.

Fig 11.

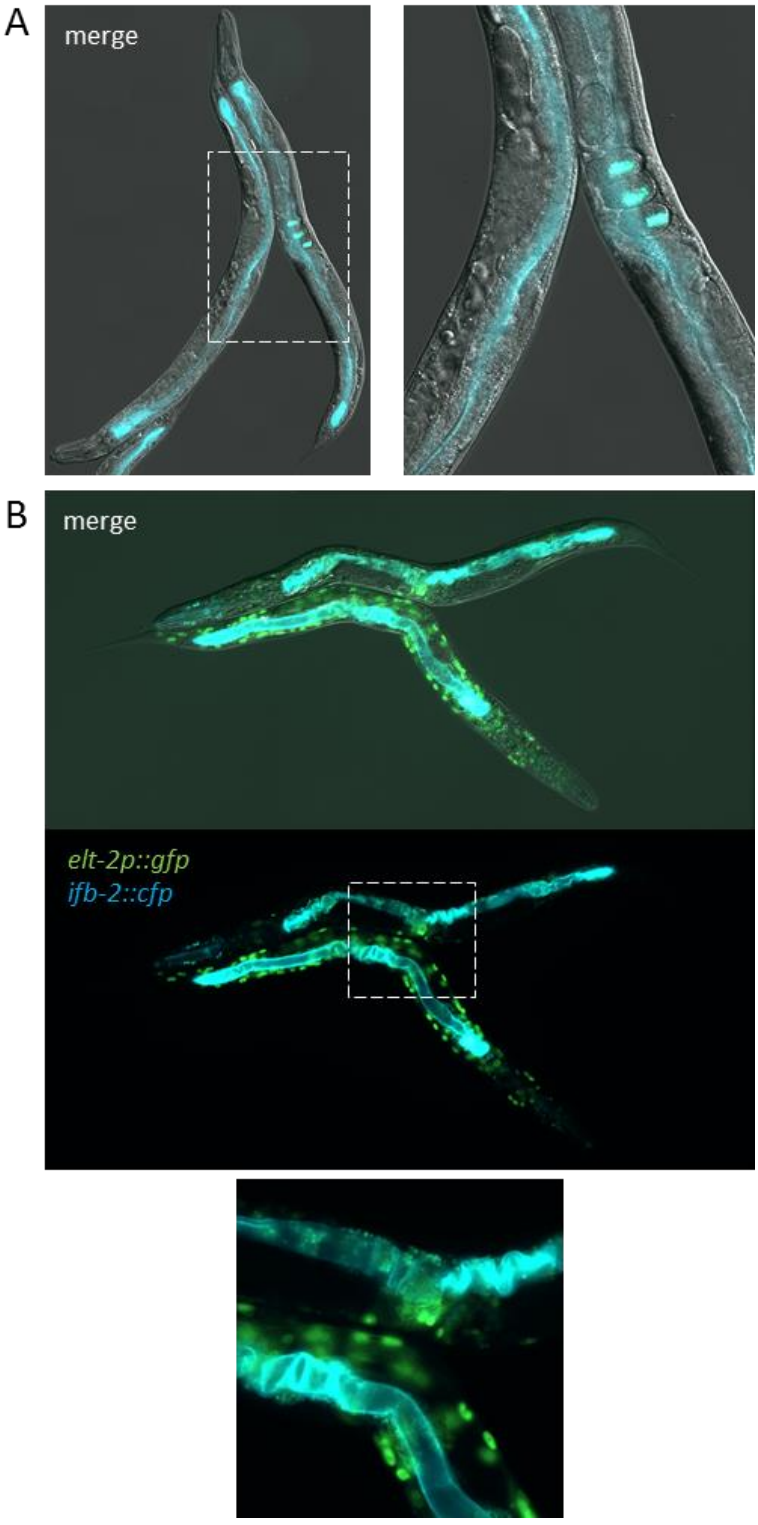


Fig 11. Knockdown of *smo-1*/SUMO can abrogate reprogramming of the somatic gonad into intestine.

(A) Worms fed from L1 on *smo-1(RNAi)* develop into sterile adults with moderate to severe gonad defects. (B) 24 hours after *hsp::elt-7* induction of *smo-1(RNAi)* treated worms, widespread ectopic *elt-2p::gfp* expression is observed as well as reprogramming in the pharynx but not in the somatic gonad. Inset: proximal gonad regions have no detectable IFB-2::CFP suggesting a failure in transorganogenesis (compare to Fig 4B).

Fig 12.

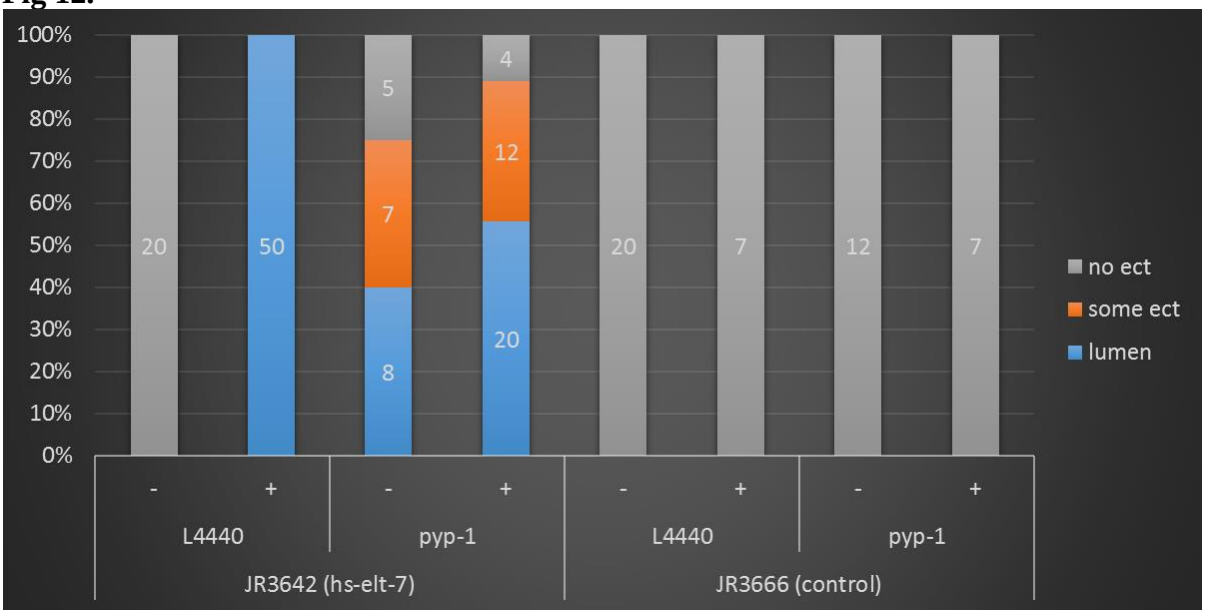

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
20
50
5
7
8
4
12
20
20
7
12
7
no ect
some ect
lumen
-
+
-
+
-
+
-
+
L4440
pyp-1
L4440
pyp-1
JR3642 (hs-elt-7)
JR3666 (control)

Fig 12. *pyp-1(RNAi)* reduces ectopic ELT-7 induced reprogramming but causes gonad reprogramming without heat shock.

Knockdown of the *pyp-1* (inorganic PYroPhosphatase) gene reproducibly inhibits somatic gonad-to-intestine transorganogenesis following ectopic *elt-7* expression, reducing the proportion of worms with a well-formed ectopic lumen by 44% (n = 36). *pyp-1(RNAi)* treated worms also appear to undergo somatic gonad-to-intestine transorganogenesis without being heat shocked 40% (n = 20) of the time. This heat shock independent reprogramming does not occur in *pyp-1(RNAi)* treated worms that do not carry the *hsp::elt-7* transgene, suggesting that knockdown of inorganic pyrophosphatase can lead to aberrant activation of the heat shock response in a tissue specific manner.

Fig 13.

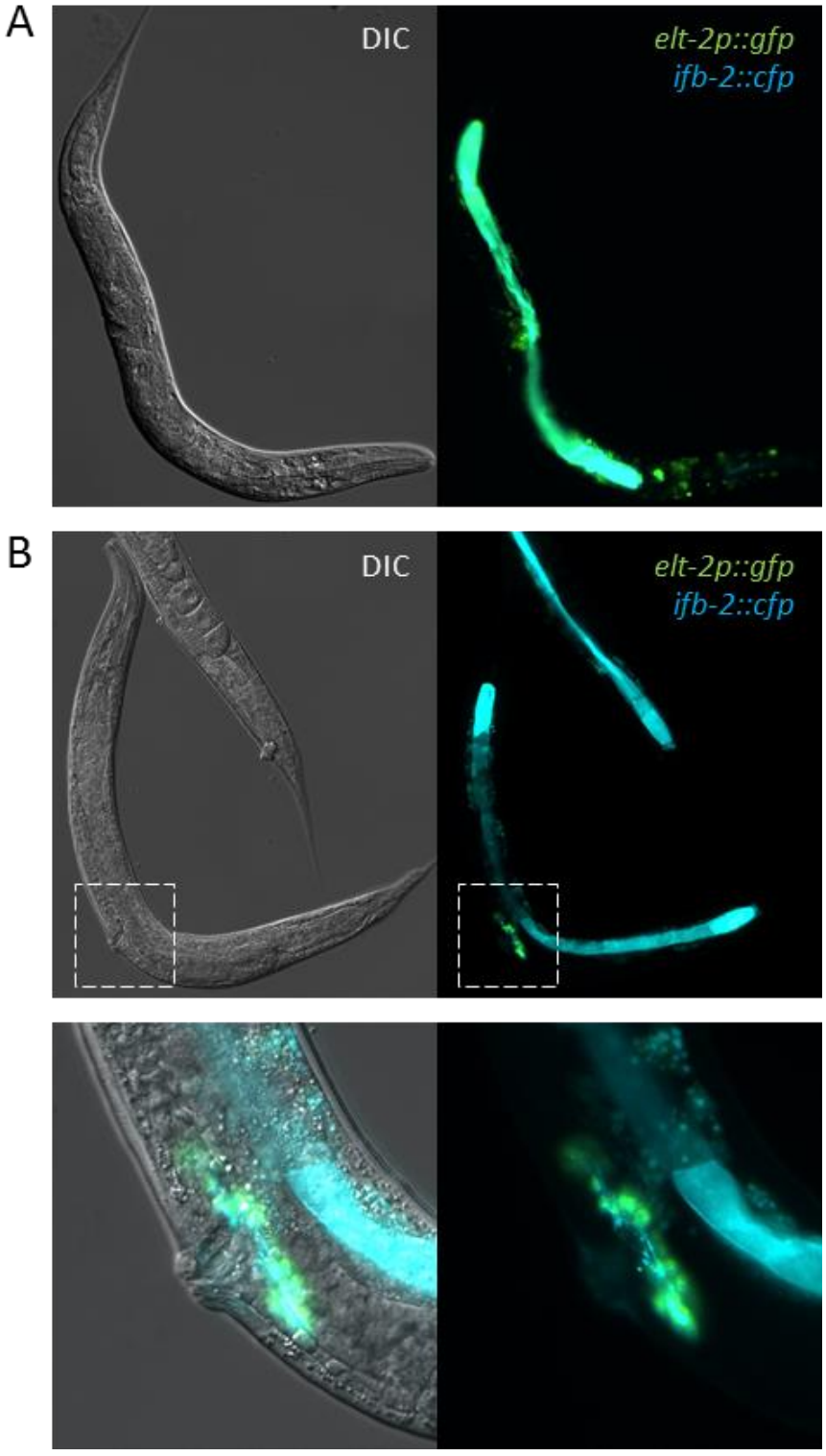


Fig 13. *pyp-1(RNAi)* results in reduced reprogramming following heat shock as well as somatic gonad reprogramming without heat shock.

(A) Representative DIC and fluorescence image of a worm fed from L1 on *pyp-1(RNAi)* for two days, heat shocked, and scored after 24 hours. Little to no ectopic IFB-2::CFP is detectable in the somatic gonad of about half of the worms. Larval development is slower compared to controls and animals display moderate gonad defects. (B) Worm fed on *pyp-1(RNAi)* for approximately three days with continuous incubation at 20^0C. Ectopic *elt-2p::gfp* and *ifb-2::cfp* expression is clearly observed in the proximal somatic gonad but not elsewhere. Inset: The uterus appears to undergo transorganogenesis in non-heat shocked worms, although the resulting ectopic intestine-like structure is smaller and not as well formed as in controls.

Fig S1.

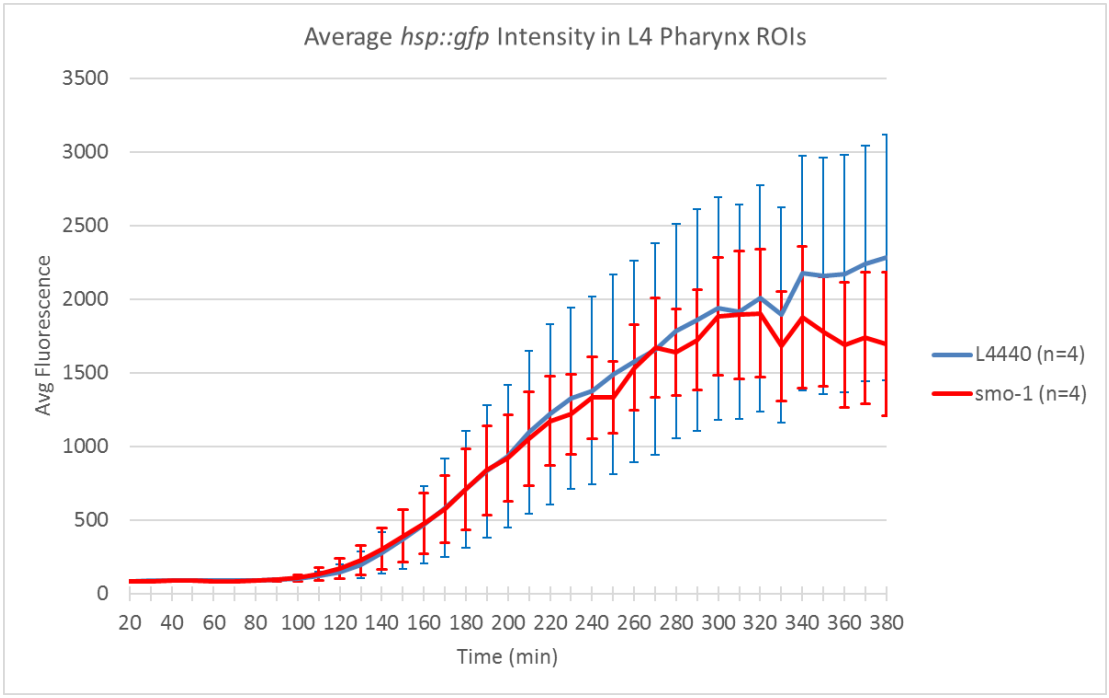

Average hsp::gfp Intensity in L4 Pharynx ROIs
Avg Fluorescence
3500
3000
2500
2000
1500
1000
500
0
20 40 60 80 100 120 140 160 180 200 220 240 260 280 300 320 340 360 380
Time (min)
L4440 (n=4)
smo-1 (n=4)

Fig S1. Average *hsp::gfp* intensity in L4 pharynx ROIs.

Time lapse imaging was used to capture the dynamics of GFP expression in the pharynx of worms treated with either L4440 control or *smo-1* RNAi for 6 hours following heat shock induced expression of a *hsp::gfp* transgene. Average fluorescence intensity is plotted, error bars show standard deviation.

Chapter Three

Temporal transcriptional profiling of transorganogenesis in *Caenorhabditis elegans* reveals broad cellular effects including activation of an intracellular pathogen-like innate immune response

Abstract

Cellular reprogramming holds promise for advancing the field of regenerative medicine. A deeper understanding of cellular reprogramming events *in vivo* and tissue specific consequences of cell fate transitions can be elucidated through research in the nematode *C. elegans*. Ectopic expression of the sole transcription factor ELT-7 can induce transorganogenesis, whereby the somatic gonad is reprogrammed into a well-structured intestine-like organ, as well as transdifferentiation of the pharynx into intestine-like tissue. The molecular mechanisms underlying the capability of a single factor to trigger the reprograming of an entire organ *in vivo* are largely unknown. To identify the biological processes involved in such transorganogenesis, we performed time course mRNA-sequencing. Analysis of temporal transcriptional dynamics and differential gene expression patterns reflects the changes in cell fate of the pharynx and somatic gonad, and indicates that tissue specific gene expression is transiently repressed in multiple non-reprogrammed cell types. Gene ontology (GO) terms related to innate immunity are enriched among upregulated genes, and an intracellular pathogen (IPR)-like transcriptional response is strongly and transiently activated following ectopic ELT-7 expression. The IPR is activated in the same tissues that are reprogrammed, leading us to propose that the capability to initiate the IPR defines a cellular context that is permissible to reprogramming to intestine by ELT-7. The convergence of innate immunity and transdifferentiation has been observed in other reprogramming events and has been described as "transflammation." Correlation of differential gene expression with transcription factor binding profiles identifies a unique cluster of five factors that are strong candidates for regulating transorganogenesis. This study

presents the first gene expression map of *in vivo* reprogramming of an entire organ and characterizes the dynamic biological processes involved.

Introduction

Establishment of cell fate during development of multicellular organisms is naturally a unidirectional process but can be experimentally reverted to a dedifferentiated or pluripotent state, or directly reprogrammed to a different fully differentiated state (Takahashi and Yamanaka, 2016; Srivastava and DeWitt, 2016). A growing repertoire of cell types can now be generated through a wide variety of reprogramming processes, ranging from the use of ectopically expressed transcription factors to exogenous pharmacological cocktails in both *in vitro* and *in vivo* settings (Esteban and Wang, 2017; Taguchi and Yamada, 2017; Vierbuchen and Wernig, 2012). Directed reprogramming of cells *in vivo* is a pioneering approach to regenerative medicine and has the potential to allow for direct repair of damaged or diseased tissues and organs (Ebrahimi, 2017; Merrell and Stanger, 2016; Ofenbauer and Tursun, 2019). For example, a degenerative disease such as liver fibrosis could be treated by *in vivo* reprogramming of myofibroblasts directly into hepatocyte-like cells (Rezvani et al., 2016; Taguchi and Yamada, 2017). Similarly, an organ injury such as a heart attack could be restored through *in vivo* reprogramming of non-functional fibroblasts into functional cardiomyocytes (Ebrahimi, 2017; Kurotsu et al., 2017; Qian et al., 2012). However, safe and effective translation of *in vivo* reprogramming therapy to the clinic requires a deeper understanding of the molecular mechanisms involved in the turnover of cell fates (Doppler et al., 2015; Smith et al., 2017).

The model research organism *Caenorhabditis elegans* provides an excellent system for studying cellular reprogramming *in vivo* (Becker and Jarriault, 2016; Hajduskova et al., 2012; Rothman and Jarriault, 2019; Spickard et al., 2018). *C. elegans* has an invariant cell lineage (Deppe et al., 1978; Sulston and Horvitz, 1977; Sulston et al., 1983) and while adult hermaphrodites possess only 959 somatic cells, they form many distinct tissues and organs including an intestine, muscles, hypodermis (skin), and a central nervous system, allowing for precise determination of starting and ending cell types involved in reprogramming events (Jarriault et al., 2008; Riddle et al., 2013; Tursun et al., 2011). The *C. elegans* proteome is also well conserved, with approximately one-third of its proteins having homologs in humans (Kim et al., 2018; Shaye and Greenwald, 2011). Therefore, *in vivo* reprogramming discoveries in the worm may likely translate to other organisms.

Early *C. elegans* embryonic cells are rapidly specified for precise specific fates yet are largely plastic and can be reprogrammed into cell types of all germ layers via ectopic expression of lineage determining factors (Fukushige and Krause, 2005; Gilleard and McGhee, 2001; Kalb et al., 1998; Zhu et al., 1998). Part way through embryogenesis, cells undergo a multipotency-to-commitment transition (MCT), after which this initial plasticity becomes restricted (Custer et al., 2014; Djabrayan et al., 2012; Spickard et al., 2018; Yuzyuk et al., 2009). Post-MCT germline cells can be induced to form teratomas through the removal of translational or chromatin repressors (Ciosk et al., 2006; Robert et al., 2014), and can also be reprogrammed into neurons through simultaneous expression of the CHE-1 transcription factor and inhibition of either LIN-53/RBBP4, Polycomb repressive complex 2, MRG-1/MRG15, or histone chaperone FACT (facilitates chromatin transcription) (Hajduskova et al., 2018; Kolundzic et al., 2018; Patel et al., 2012; Seelk et al., 2016; Tursun et al., 2011).

Work in our lab has demonstrated that ectopic expression of the endoderm directing GATA-type transcription factor ELT-7 induces the fully formed and functional pharynx to transdifferentiate into intestine-like tissue and, remarkably, the developing somatic gonad is reprogrammed and restructured into a secondary intestine-like organ, a process we have termed "transorganogenesis" (Riddle et al., 2013; Riddle et al., 2016).

Complete reprogramming of an organ *in vivo* driven by a single factor is, to our knowledge, an unprecedented phenomenon. The pharynx and somatic gonad are the only cell types that appear to be susceptible to reprogramming by ELT-7. This tissue specificity may be attributable to the developmental activity of the forkhead transcription factor PHA-4, which is integral for the development of the pharynx, as well as being involved in intestine and gonad development (Anokye-Danso et al., 2008; Chen and Riddle, 2008; Gaudet and Mango, 2002; Horner et al., 1998; Kalb et al., 1998; Mango et al., 1994; Riddle et al., 2016). In *pha-4(-)* larvae, the pharynx is not properly differentiated and does not become reprogrammed after ectopic ELT-7 expression, but PHA-4 activity is not required during the reprogramming process (Riddle et al., 2016). The mechanisms by which cell fate conversion is initiated and potentiated are unknown and additional regulators of cellular reprogramming in this context have not been identified. Furthermore, while transmission electron microscopy illustrates the dramatic ultrastructural remodeling of pharyngeal and gonadal tissues (Riddle et al., 2013; Riddle et al., 2016), ectopic ELT-7 is broadly expressed, and additional cellular, organ, or organismal consequences of *in vivo* forced reprogramming have not been thoroughly characterized.

To identify the biological processes that are dynamically modulated throughout the reprogramming process, we performed temporal transcriptional profiling of *C. elegans* at 5

time points after ectopic ELT-7 expression and quantified gene expression changes for 20,094 protein coding genes, as well as for a time-matched control strain. Here we present the first organism-wide gene expression map of transorganogenesis. We highlight gene expression profiles that capture the changes in cell fate of the pharynx and somatic gonad. In general, intestine-specific genes which may play important roles during reprogramming are upregulated, while structural and developmental genes in the pharynx and somatic gonad are downregulated. Spatial analysis of gene expression performed *in silico* using previously defined tissue enriched gene expression patterns (Spencer et al., 2011) shows that non-reprogrammed tissues vary in their transcriptional response to ELT-7 in two primary ways. First, gene expression for some tissues is relatively unaffected on average. Second, other tissues show transiently and specifically repressed gene expression resulting from ectopic ELT-7 activity. We next examined the enrichment of gene ontologies within clusters of upregulated and downregulated genes with unique temporal differential expression patterns. This analysis revealed that gene ontology terms relating to innate immunity processes are particularly strongly enriched. Further, gene set enrichment analysis revealed that genes relating to an intracellular pathogen response (IPR) are strongly and transiently upregulated following ectopic ELT-7 expression, and spatial analysis of IPR activation indicates that the intestine, pharynx, and somatic gonad are specifically able to initiate this transcriptional response. Lastly, we compare transcription factor binding site (TFBS) profiles (Kudron et al., 2017) with our differentially expressed gene (DEG) sets and identified a group of five factors, ELT-2, NHR-80, NHR-28, FOS-1, and PQM-1, which correlate with upregulated DEGs and represent potential regulators of transorganogenesis. In light of our findings, we propose that successful ELT-7 directed cellular reprogramming is dependent on a cellular

context that has both (1) been defined by developmental PHA-4 activity, and (2) is capable of activating an intracellular pathogen-like transcriptional response.

Results

Transcriptional profiling of in vivo *cellular reprogramming and transorganogenesis using mRNA sequencing*

Previous work revealed that the ectopic overexpression of the intestine-specific GATA factor gene *elt-7* is sufficient to induce direct *in vivo* transdifferentiation of the fully differentiated pharynx, and transorganogenesis of the developing somatic gonad, including the hermaphrodite uterus and spermathecae, and the male vas deferens, into intestine-like tissue (Riddle et al., 2013; Riddle et al., 2016). Compelling evidence from transmission electron microscopy of fully reprogrammed tissues, differential interference contrast microscopy, and fluorescent reporter expression suggests that the pharynx and somatic gonad undergo complete and direct cell fate conversion (Fig 1A). Such dirastic changes in cell types is likely to involve significant transcriptional changes.

We aimed to characterize these transcriptional changes throughout the time-course of animals undergoing forced *in vivo* tissue and whole-organ reprogramming and to identify key gene expression signatures to gain a clearer picture of the process of transorganogenesis. Our approach was to utilize high throughput sequencing of RNA (RNA-seq) purified from synchronized populations of worms at four time points following heat shock of a strain carrying the integrated, multicopy trangenic arrays wIs125[*hsp-16.2p::elt-7 + hsp-16.41p::elt-7 + rol-6(su1006)*]; rrIs1[*elt-2p::gfp*]; kcIs6[*ifb-2::cfp*], hereafter referred to as

hsp::elt-7. All worms were heat shocked at the early L4 stage, which has previously been shown to be the point in development that results in the most penetrant transorganogenesis phenotype (Riddle et al., 2016). In addition to a t = 0 hours post heat shock (hrs PHS) reference sample, collected immediately before heat shock treatment, total RNA was extracted from samples at 3, 6, 12, and 30 hrs PHS (Fig 1B). These times were chosen as they correspond to distinct stages of the reprogramming process as determined by observation of the *elt-2p::gfp* transcriptional reporter for early intestine fate and the *ifb-2::cfp* translational reporter for intestine differentiation (Fig 1C). Widespread expression of *elt-2p::gfp* is seen at 3 hrs PHS, suggesting that at this early time point in the reprogramming process sufficient functional ELT-7 protein has been produced to begin to drive transcriptional changes. Expression of *ifb-2::cfp* can first be seen in the pharynx at 6 hrs PHS, and in the somatic gonad at 12 hrs PHS, indicating that these tissues have begun to acquire an intestine-like cell fate. By 30 hrs PHS, the pharynx and somatic gonad appear to be fully reprogrammed with *ifb-2::cfp* forming a striking secondary intestine-like lumen structure. In order to account for any transcriptional changes resulting from heat shock treatment or the *rol-6(su1006)* marker, and to monitor induction of the heat shock response, a control strain with the integrated, multicopy transgene dvIs70[*hsp-16.2p::gfp + rol-6(su1006)*], hereafter referred to as *hsp::gfp*, was treated in parallel to the *hsp::elt-7* strain.

Total RNA was poly-A purified to build mRNA libraries, which were sequenced using an Illumina NextSeq 500 instrument. An average of 50 ± 18.5 million reads per sample were obtained for single end runs and an average of 23 ± 1.2 million reads per sample for paired-end runs (Table 1). Read processing and QC was performed with the web-based Galaxy bioinformatics tools using the usegalaxy.org public server (Afgan et al., 2018; Blankenberg

et al., 2014). Read pseudoalignment and quantification of gene expression levels was performed using Salmon (Patro et al., 2017) with the WS263 reference *C. elegans* transcriptome, which has annotations for 20,094 protein coding genes (Fig 1D).

Exploratory data analysis of transcriptional profiles using principle components (PC) clearly differentiates two temporal transcriptional paths (Fig 1E). The control samples are separated along the PC2 axis in temporal order, indicating that this PC correlates with normal developmental gene expression changes occuring from L4 to adult stage worms. The *hsp::elt-7* samples are separated from the controls by PC1 and follow a steep curved path indicating that there are numerous genes with high variance and transient expression profiles. *hsp::elt-7* samples also follow in temporal order along the curved path; however, they are closer together along PC2 and the 30 hrs PHS *hsp::elt-7* results show a PC2 coordinate that is placed between the 6 and 12 hrs PHS for *hsp::gfp*. These data are in agreement with the observed developmental arrest phenotype that accompanies transorganogenesis; heat shocked L4 stage *hsp::elt-7* worms do not develop into adults or produce offspring. PC1 and PC2 together explain up to 82% of the gene expression variance among samples, indicating that the majority of gene expression changes observed result from the experimental treatment and design. Furthermore, there is close agreement between biological replicates (Fig S1), suggesting that on average the process of transorganogenesis is highly reproducible at the transcriptional level.

Statistical analysis of differential gene expression between time-point matched *hsp::elt-7* versus *hsp::gfp* samples was carried out using DESeq2 (Love et al., 2014). Using a false discovery rate of 0.01 for each time point, we detected between 3,729 – 4,526 genes that were significantly upregulated compared to controls, and between 3,054 – 6,348 genes that

were significantly downregulated compared to controls (Fig 1F). At 3 and 6 hrs PHS, more genes are upregulated than downregulated, while at 12 and 30 hrs PHS, more genes are downregulated than upregulated. Differential gene expression fold change values range from over $\pm 2^{10}$ for all time points, except for 3 hrs PHS where genes are not as extremely downregulated (Fig S2). These data comprise the first temporal gene expression map of *in vivo* whole-organ reprogramming and the basis for futher insight into the effects of *elt-7* directed transdifferentiation.

Gene expression profiles capture the changes in cell fate of the pharynx and somatic gonad following ELT-7 *ectopic expression*

With the complete transcriptional profiles for all protein coding genes in the genome, we first looked at the dynamics of gene expression changes for factors related to the structure and function of the tissue types being converted, specifically the intestine, pharynx, and uterus (Fig 2). We found that many intestine factors show significantly increased expression following *hsp::elt-7* induction. The two GATA factor paralogs *elt-7* and *elt-2* are critical intestine differentiation transcription factors that are normally transcribed starting mid-embryogenesis and continuously throughout larval development and adulthood, functioning to activate the transcription of thousands of intestine genes (Dineen et al., 2018; Fukushige et al., 1998; Maduro, 2015; McGhee et al., 2009; Sommermann et al., 2010). Normalized counts of RNA-seq reads show a 4.9-fold increase in *elt-2* mRNA at 3 hrs PHS followed by a leveling off to approximately twice baseline expression (Fig 2A). This pattern may reflect initial wide-spread activation of *elt-2* by ectopic *elt-7* that later becomes restricted to only the tissues that undergo transdifferentiation and agrees with the observed spatio-temporal

expression pattern for the integrated *elt-2p::gfp* transcriptional reporter (Fig 1C and (Riddle et al., 2016)). *elt-7* shows a continuous increase in expression up to 13-fold at 12 hrs PHS and then subsequently drops off to 1.7-fold greater than control level by 30 hrs PHS. The majority of *elt-7* mRNA detected likely originates from the multicopy *hsp::elt-7* transgenic array and provides insight into the strength and perdurance of transcription from the heat shock promoter in this system.

During normal embryonic development, the redundantly acting GATA factors *end-1* and *end-3* function upstream of *elt-7* and *elt-2* in the specification of intestinal cells as part of a feed-forward transcription factor cascade (Boeck et al., 2011; Maduro and Rothman, 2002; Maduro et al., 2005; Zhu et al., 1997; Zhu et al., 1998) (Chapter 2; Fig 1). *end-1* and *end-3* are expressed only transiently within the early E-cell lineage (Raj et al., 2010); however, we found that *end-1* is reactivated post-embryonically in L4 stage worms after *hsp::elt-7* induction with a 14-fold increase in expression at 3 hrs PHS, and *end-3* shows a 3.7-fold increase in expression at 6 hrs PHS (Fig 2A), although it should be noted that these differences were not deemed significant at our chosen false discovery rate of 0.01. A few counts are also detected in controls between 0 and 12 hrs PHS and then clearly expressed at 30 hrs PHS, most likely due to normal embryonic development, as the controls contain adult hermaphrodites with fertilized embryos. The intermediate filament protein IFB-2 is expressed specifically in the intestine from late embryogenesis through adulthood and contributes to the cytoskeletal structure of the terminal web of the intestinal lumen (Bossinger et al., 2004; Dodemont et al., 1994). *ifb-2* expression following *hsp::elt-7* induction mirrors the pattern of *elt-2* with a 4-fold increase over controls at 3 hrs PHS. Similar to the observed *elt-7* expression, a significant proportion of the transcripts may

originate from the integrated *ifb-2::cfp* translational fusion transgene present in the strain (Fig 2A). However, unlike the rapid, widespread appearance of the *elt-2p::gfp* reporter, a burst in *ifb-2::cfp* signal is not observed and IFB-2::CFP protein appears only later within the pharynx and somatic gonad. This suggests that although *ifb-2* is transcribed early, its translation may be inhibited or the ectopic protein product may be rapidly degraded initially.

The bZip transcription factor SKN-1 (the mammalian Nrf ortholog) has two canonical functions in *C. elegans*: (1) as a maternal factor required for specification of the EMS blastomere during early embryogenesis, and (2) post-embryonically as a regulator of the oxidative stress response pathway (the phylogenetically conserved function), and is expressed strongly in the intestine as well as the hypodermis and ASI neurons (An and Blackwell, 2003; Bowerman et al., 1992). Expression of *skn-1* in controls increases at later time points correlating with development into adulthood and its function as a maternally provided factor. Following ectopic *elt-7* expression, *skn-1* mRNA increases 2.2-fold by 6 hrs PHS and then decreases again. As the *hsp::elt-7* strain undergoes developmental arrest, this increase in *skn-1* is likely not attributable to its maternal function but rather its role in cellular response to stress, suggesting that cellular reprogramming is inducing cell stress that is independent of the heat shock treatment.

We further confirmed that important structural and regulatory factors in pharyngeal tissue show decreased gene expression during forced transdifferentiation into intestine. It was previously observed that a reporter for the pharynx muscle-specific myosin gene *myo-2* is lost in pharynx cells that undergo transdifferentiation (Okkema et al., 1993; Riddle et al., 2013). Both strains initially have a parallel drop in *myo-2* expression from 0 to 3 hrs PHS, which may be attributable to either developmental or heat shock specific effects on

transcription. In the *hsp::elt-7* strain, *myo-2* expression is decreased 30-fold compared to controls at 12 hrs PHS (Fig 2B). Similarly, the intermediate filament protein *ifa-1* is primarily expressed in the pharynx and also shows decreased expression, albeit not as strongly as *myo-2*: it is downregulated 2.2- to 2.5-fold from 3 to 12 hrs PHS. It should be noted that *ifa-1* is also expressed in the vulva and rectum, which may account for the remaining transcripts detected (Dodemont et al., 1994; Karabinos et al., 2001; Karabinos et al., 2003). The *C. elegans* homeobox and NK-2 family homeodomain factor CEH-22 is required for normal pharynx development and activates pharyngeal muscle gene expression (Okkema and Fire, 1994; Okkema et al., 1997). *ceh-22* mRNA is rapidly diminished following *hsp::elt-7* induction with an 11-fold decrease at 6 hrs PHS, suggesting a collapse of the pharyngeal gene regulatory network. This loss of *ceh-22* activity may contribute to efficient cellular reprogramming of the pharynx into intestine.

Gene expression profiles also capture the loss of key uterine factors during transorganogenesis (Fig 2C). *ule-3* is representative of a class of genes that were identified as being uniquely expressed in the uterine lumen (Zimmerman et al., 2015). Control *ule-3* gene expression shows strong activation corresponding to the onset of adulthood and maturation of the uterus. No activation of *ule-3* expression is observed after ectopic *elt-7,* suggesting the complete loss of uterine organ identity. The Ras and EF-hand domain containing homolog *rsef-1* is expressed in uterine lumen epithelial cells and mediates proper gonad development and outgrowth of the uterine seam cell in late larval development (Ghosh and Sternberg, 2014). *rsef-1* gene expression is comparable to controls at 0 and 3 hrs PHS, but then significantly deceases in the *hsp::elt-7* strain while increasing in the *hsp::gfp* strain, resulting in roughly 4-fold downregulation from 6 to 12 hrs PHS. Lastly, we examined expression of

the gonad-specific forkhead transcription factor *fkh-6*, which is required for proper gonadogenesis (Chang et al., 2004; Hope et al., 2003). *fkh-6* expression is gradually reduced in controls but is rapidly downregulated by 3-fold 3 hours following ectopic *elt-7* expression.

These results highlight the changes in cell fate that occur following widespread *elt-7* expression and provide additional molecular evidence for transdifferentiation of the pharynx and transorganogenesis of the uterus into intestine tissues. Overall, genes known to be important for the development and differentiation of the intestine have increased expression levels, and genes that are similarly important for the development and functioning of the pharynx and uterus have decreased expression levels. Furthermore, each sample represents the average gene expression present in a large population of worms so these changes, particularly the near complete loss in expression of genes such as *myo-2* and *ule-3*, are occurring in highly reproducible patterns.

Different tissues vary in their transcriptional responses to ectopic elt-7

The pharynx and somatic gonad are the only tissues that undergo complete cellular reprogramming following ectopic *elt-7* expression. However, the *hsp::elt-7* transgene is expressed ubiquitously following heat shock treatment and transcriptionally active ELT-7 protein is produced, as evidenced by *elt-2p::gfp* expression throughout the animal. In addition to the transorganogenesis and transdifferentiation phenotypes, it has also been observed that these worms experience chemotaxis defects, molting defects, reduced movement, and developmental arrest (Riddle et al., 2016). Based on these phenotypes, we hypothesized that significant transcriptional changes may be occurring in other tissue types as well.

Our transcriptional profiling data was generated from populations of whole worms and therefore contains information on gene expression dynamics in all tissue types. A benefit to using *C. elegans* as a model system is the deep knowledge of individual cell and tissue types and their corresponding gene expression patterns. In order to query how different tissues respond transcriptionally to ectopic *elt-7,* we compared our RNA-seq results to selectively enriched gene sets from a previous study that detailed the spatio-temporal transcriptome of *C. elegans* (Spencer et al., 2011). We observed overall trends in gene expression for all genes within a given tissue-enriched gene set and plotted the fold-changes in gene expression (Fig 3). All tissue-enriched gene sets contain some genes that are more than 2-fold up- or downregulated. When we calculated the average fold change for all genes specifically expressed in each tissue we found two apparent trends. First, some tissues show relatively unchanged average gene expression following *hsp::elt-7* induction compared to the *hsp::gfp* control, and second, a subtle but clear trend in multiple tissues, where average gene expression levels decrease from 3 to 6 hrs PHS and subsequently increase to levels equal or even greater than controls. However, the average fold change values could potentially be skewed by strong outliers and may not accurately describe the actual trends for the majority of the tissue-type expressed genes. We then tested whether each tissue-specific gene set significantly correlates with up- and downregulated differentially expressed genes at each time point using Fisher's exact test (Fig 3, Fig S3). The results further confirmed and support the unexpected trends in gene expression for different tissue types.

Overall gene expression changes vary up to a maximum of over 1,000-fold up- and down-regulated, but on average there is less than a 1.5-fold increase in expression across all time points. In tissues displaying the unchanged average gene expression trend, such as A-

class neurons, we similarly found weak or no statistical correlations with DEGs. In hypodermal tissue there is a 1.43-fold decrease in tissue-specific gene expression at 6 hrs PHS on average, but there is a strong correlation with downregulated DEGs (p-value = 1.99×10^{-18}). At 12 hrs PHS there is a 1.15-fold increase in average hypodermal gene expression that correlates with upregulated DEGs (p-value 1.01×10^{-9}). Body muscle-specific gene expression also displays the transient downregulation trend with a 1.82-fold decrease in average expression at 6 hrs PHS that very strongly correlated with downregulated DEGs (p-value = 3.44×10^{-52}).

Unexpectedly, pharynx muscle-specific genes also appear to show only transiently downregulated expression. This could be partially explained by the fact that not all of the cells in the pharynx undergo transdifferentiation in every worm – complete reprogramming of the entire pharynx is only rarely observed. Additionally, some of the genes may not be exclusively expressed in the pharynx muscle and our analysis is unable to define spatial expression for genes that are expressed broadly or ectopically. We also observed unexpected results for intestine-specific genes that show a more wave-like pattern in average expression. Intestine genes are correlated with downregulated DEGs at all time points and correlated with upregulated genes only at 12 hrs PHS (p-value = 2.81×10^{-11}). This may be the result of complex gene signatures originating from the endogenous intestine as well as the ectopically forming intestine tissues in the pharynx and somatic gonad. The vast majority of intestine-specific gene expression is likely from the endogenous intestine and the normal endoderm gene regulatory network may respond to overexpression of *elt-7* with compensatory negative feedback, resulting in the downregulation of genes observed.

Comparison of the overall correlations between tissue type enriched gene sets and DEGs reveals a greater frequency of correlation with downregulated DEGs than with upregulated DEGs, particularly at 3 and 6 hrs PHS (Fig S3). This is interesting given that there are more upregulated than downregulated DEGs at 3 and 6 hrs PHS (Fig 1F) and that gene expression is globally increased on average. This suggests that certain cell type-specific gene regulatory networks are transiently disrupted in response to ELT-7 but are gradually able to recover and to avoid becoming reprogrammed.

A broad array of cellular functions are enriched in gene ontology of differentially expressed genes

Thousands of genes are significantly differentially expressed after ectopic *elt-7* expression. A useful strategy for gaining insight into the cellular functions and biochemical pathways that may be affected throughout the *in vivo* cellular reprogramming processes is to look for enrichment of gene ontologies (GO) within identified gene sets. The time course transcriptomic data enables a search for temporal patterns in cellular perturbations. To enrich for GO terms across many distinct clusters of genes simultaneously, we used the clusterProfiler Bioconductor software package (Yu et al., 2012). We initially performed GO enrichment on eight DEG clusters defined by the experimental design, with one upregulated and one downregulated cluster for each time-point. However, this yielded strong but highly redundant GO enrichment profiles, most likely owing to the large proportion of genes that are differentially expressed at multiple time-points. We therefore chose to cluster upregulated and downregulated DEGs according to the particular set of time points at which they are differentially expressed (Fig 4A-B). The resulting 30 DEG clusters – designated U01-U15

for upregulated gene clusters and D01-D15 for downregulated gene clusters – allowed for GO enrichment among groups of genes with unique expression paths following *hsp::elt-7* induction and minimized the effects of genes that are up- or downregulated at multiple time-points on redundancy of GO terms.

We found a diverse assortment of GO terms enriched across clusters (Fig 4C-D, Fig S4-S9). For the 15 upregulated DEG clusters, we identified a total of: 575 biological process, 157 cellular component, and 149 molecular function terms (FDR < 0.05). We also searched the Reactome pathway database and found 171 enriched pathways (Fabregat et al., 2017). Among the 15 downregulated DEG clusters, we identified: 1,034 biological process, 294 cellular component, 230 molecular function, and 201 Reactome pathway terms (FDR < 0.05). For upregulated DEG clusters, the most significantly enriched biological process term is GO:0008360 "regulation of cell shape" in cluster U15 (p-value = 7.6×10^{-37}) (Fig 4C). For cellular component, molecular function, and Reactome pathway the most significantly enriched terms are the related GO:0022626 "cytosolic ribosome" (p-value = 5.2×10^{-45}) (Fig S4), GO:0003735 "structural constituent of ribosome" (p-value = 6.0×10^{-39}) (Fig S5), and R-CEL-72706 "GTP hydrolysis and joining of the 60S ribosomal subunit" (p-value = 2.8×10^{-52}) (Fig S6) respectively, found in cluster U01 (Up.3). The most significantly enriched biological process, cellular component, and molecular function GO terms for downregulated DEGs are associated with cluster D14 (Down.12.30); GO:0022402 "cell cycle process" (p-value = 6.6×10^{-58}) (Fig 4D), GO:0005694 "chromosome" (p-value = 1.9×10^{-56}) (Fig S7), and GO:0003723 "RNA binding" (p-value = 1.6×10^{-25}) (Fig S8), respectively. The most significantly enriched Reactome pathway is R-CEL-112316 "Neuronal System" (p-value = 7.7×10^{-12}) in cluster D02 (Down.3.6) (Fig S9).

Dotplots of the top up- and downregulated GO terms illustrate the relative temporal sequence of cellular perturbations following *hsp::elt-7* induction (Fig 4C-D, Fig S4-S9). The 30 DEG clusters are largely associated with unique gene ontologies. Overall, 79.3% (n = 1,326) of the upregulated GO terms and 72.3% (n = 2,432) of the downregulated GO terms were associated with only a single DEG cluster. This suggests that cellular processes are generally not differentially or repeatedly perturbed by *elt-7* activity, rather there are reproducible effects that occur during particular time spans in the course of *in vivo* reprogramming. A notable exception to this trend is seen for biological process GO terms related to the *C. elegans* immune response, which are significantly enriched throughout the time course, appearing in four clusters, including U08 (Up.3.6.12.30), which contains 478 genes upregulated at all time points post heat shock (Fig 4C).

Ectopic elt-7 expression induces an intracellular pathogen-like transcriptional innate immune response

In addition to GO annotations, the *C. elegans* knowledgebase contains thousands of publicly available data sets from gene expression studies involving a wide array of experimental conditions. We asked whether there are highly variably expressed genes that we had detected that correlate with gene expression changes previously found to be associated with a particular mutation, stress, or other condition. Initially, we were most interested in identifying a pattern of gene expression changes related to the first principle component of our PCA results (Fig 1E). We extracted a list of the top 225 genes with the greatest contribution to PC1 and cross compared this gene set to over 2,000 other *C. elegans* experimental gene expression data sets using the web-based WormExp application (Yang et

al., 2016). Four out of the top ten enrichment results showed highly significant associations with gene sets originating from the same study (p-values $\leq 3.33 \times 10^{-40}$), which profiled the temporal transcriptional response in *C. elegans* to intracellular infection by the microsporidian pathogen *Nematocida parisii* (Bakowski et al., 2014).

Microsporidia are a group of unicellular spore-forming intracellular pathogens related to fungi (Keeling and Fast, 2002). The *N. parisii* species is known to infect the intestinal cells of *C. elegans* in the wild after ingestion of spores (Troemel et al., 2008). Its replicative lifecycle is approximately three days and infected worms experience larval arrest and lethal phenotypes (Luallen et al., 2015). The most strongly upregulated gene following *N. parisii* infection was *pals-5* (Protein containing ALS2cr12 signature) (Bakowski et al., 2014). *pals-5* is also up to 555-fold upregulated and dynamically expressed after ectopic expression of *elt-7* with a sharp peak in expression at 6 hrs PHS.

The Bakowski et al. study profiled the transcriptional responses to *N. parisii* at 8, 16, 30, 40, and 64 hours post infection (hpi), corresponding to distinct phases in the microsporidian lifecycle. They detected between 50-150 upregulated genes and between 4-400 downregulated genes that were significantly differentially expressed at the various time points. These DEGs do not closely resemble other types of infection or immune responses and Bakowski et al. collectively termed the observed gene expression changes the intracellular pathogen response (IPR). We looked at whether our transcriptional profiling data fit with the IPR by plotting the combined gene expression profiles for each set of DEGs from the Bakowski et al. study, along with the *hsp::elt-7* versus *hsp::gfp* $\log_2$ fold changes (Fig S10). For genes that are downregulated post *N. parisii* infection, there is a decreasing trend in expression post *hsp::elt-7* induction as well. For genes that are upregulated post

infection, we found that nearly all of them are transiently and significantly upregulated following *hsp::elt-7* induction in a pattern similar to that of *pals-5* alone. The strongest correlation is seen for genes upregulated 16 hpi (Fig 5A). IPR gene expression was either not observed or detected at very low levels in the *hsp::gfp* strain but after *hsp::elt-7* induction the IPR genes are dramatically induced starting at 3 hrs PHS, reaching an expression maximum around 6 hrs PHS, and decreasing to close to control levels by 12 or 30 hrs PHS. IPR genes upregulated 16 hpi are expressed 206.5-fold higher on average 6 hours following ectopic *elt-7* expression (Fig 5B).

We next investigated the detailed spatial expression pattern for the IPR. As a bellwether for activation of the IPR, Bakowski et al. also constructed a *pals-5p::gfp* reporter transgene and demonstrated that GFP is expressed in intestinal cells following intracellular pathogen infection, and is also induced in response to certain other stressors such as inhibition of the ubiquitin-proteasome system (Bakowski et al., 2014). As a control, we first tested the *pals-5p::gfp*-containing strain for GFP expression following our standard 30 minute, 33^0C heat shock treatment and did not observe any increase in signal (Fig 5C). We then crossed the *hsp::elt-7* strain with the *pals-5p::gfp* IPR marker strain and observed the time-course of GFP expression following heat shock induction of *elt-7*. We saw a strong increase in GFP signal within most cells of the intestine beginning around 6 hrs PHS and reaching a maximum at approximately 12 hrs PHS, which aligns with previous reports on IPR activation and our transcriptional profiling data. However, in addition to GFP expression in the intestine, we also found that *pals-5p::gfp* is activated both in the proximal somatic gonad and throughout the pharynx (Fig 5C), the same tissues that are susceptible to reprogramming into intestine.

We also tested whether ectopic expression of additional factors that are involved in transdifferentiation of tissues into intestine has an effect on IPR activation. Ectopic *end-1* expression alone has previously been shown to be insufficient to induce transdifferentiation of the pharynx or transorganogenesis of the uterus (Riddle et al., 2013; Riddle et al., 2016), but when crossed with the *pals-5p::gfp* reporter we observed a spatio-temporal GFP expression pattern similar to that of ectopic *elt-7*, albeit generally weaker in intensity (Fig 5C). Ectopic *end-3* is sufficient for cellular reprogramming into intestine and *pals-5p::gfp* expression is also observed in the somatic gonad, although it appeared less widespread and more concentrated in what appears to be the spermatheca and some cells of the vulva.

It is possible that activation of the IPR is an intestine tissue-specific feature, and expression of the *pals-5p::gfp* reporter is only observed in pharynx and somatic gonad as a consequence of these tissues being reprogrammed into intestine, thereby gaining the capability for activating the IPR. We therefore tested the effects on IPR activation for factors that are involved in controlling cell fates other than intestine. The transcription factors ELT-1 and HLH-1 can drive the acquisition of skin and muscle cell fates, respectively, of early blastomeres in embryos prior to the multipotency-to-commitment transition (MCT) (Fukushige and Krause, 2005; Gilleard and McGhee, 2001; Rothman and Jarriault, 2019; Spickard et al., 2018). After the MCT, which occurs early in embryogenesis, ectopic expression of *elt-1* or *hlh-1* does not induce any cellular reprogramming. After ectopic *elt-1* or *hlh-1* expression in worms carrying the *pals-5p::gfp* transgene, GFP is observed in intestine, pharynx, and somatic gonad tissues in a pattern similar to that observed following ectopic *end-3* (Fig 5C). These results suggest that activation of the IPR in these tissues is not dependent on transdifferentiation but rather appears to be an inherent capability. Of the

conditions we tested, this tissue-intrinsic immune response to intracellular pathogen infection is most strongly activated by *elt-7* and occurs within the same tissues that are most susceptible to forced reprogramming to an intestine cell fate.

Transcription factor biding site profiles combined with ELT-7-induced differentially expressed genes reveal novel transcription factor associations

Our differential gene expression analysis detected thousands of upregulated and downregulated genes throughout the time course of ectopic *elt-7* directed transorganogenesis (Fig 1F). These gene expression changes could arise by numerous mechanisms, including modulation in the activity of upstream transcription factors. We therefore asked whether there are other known transcription factors that are strongly affected by *elt-7* reprogramming and are able to initiate transcription of additional genes that might aid in cell fate changes or that show reduced transcription initiating activity that increases susceptibility to forced cell fate change. The recently established model organism encyclopedia of regulatory networks (modERN) resource contains genome-wide transcription factor binding site (TFBS) profiles for 217 GFP-tagged *C. elegans* transcription factors, including well studied developmental regulators as well as transcription factors with poorly characterized functions (Kudron et al., 2017). We cross-referenced the DEGs from our temporal transcriptional profiling analysis with the transcription factor activity of the 217 modERN profiled transcription factors. We mapped the genome-wide TFBS profiles to putative promoter regions, defined as the interval from 1,000 base pairs upstream to 250 base pairs downstream of the transcription start site, for all 20,094 protein coding genes, resulting in sets of genes for which potential transcriptional regulation by a given factor was inferred. The statistical likelihood that a

significant proportion of genes regulated by a given transcription factor are also differentially expressed at a particular time point following ectopic *elt-7* expression was determined using Fisher's exact test (Table S1).

Correlations between DEGs and TFBS profiles are weaker at earlier time points and stronger at later time points overall. Using Benjamini-Hochberg adjusted p-values and a false discovery rate of 0.1 (Benjamini and Hochberg, 1995), 9% of transcription factors contain significantly enriched binding among DEGs upregulated 3 hrs PHS, while 98% of transcription factors contain significantly enriched binding among DEGs upregulated 30 hrs PHS. We considered the possibility that transcription factor-driven differential gene expression at late time points may largely be the result of developmental differences between experimental and control strains and we therefore omitted the 30 hrs PHS time point from further TFBS correlation analysis. The $\log_{10}$ p-values for all DEG-TFBS correlations at 3, 6, and 12 hrs PHS were plotted and clustered by row similarity and values were scaled by the time point-specific p-value ranges to visualize the most significant correlations across time points (Fig 6).

We found that upregulated DEGs at 3, 6, and 12 hrs PHS are significantly enriched for genes that are bound by the ELT-2 transcription factor, with p-values between $7.96 \times 10^{-9} > 1.38 \times 10^{-23}$, while downregulated DEGs at 3 hrs PHS are not significantly enriched for ELT-2 binding (p-value = 0.566). ELT-2 is the primary transcription factor involved in intestine differentiation, activating thousands of intestine-specific genes, and is itself directly activated by ELT-7 (McGhee et al., 2009; Sommermann et al., 2010). This correlation was therefore to be expected for transcriptional changes associated with reprogramming of additional tissues into intestine and could be considered a positive control for our DEG-TFBS analysis. ELT-2

enrichment also serves as an anchor point for identifying other transcription factors that may function in concert with ELT-7 and ELT-2 to promote transorganogenesis.

ELT-2 makes up part of a cluster of 5 transcription factors with similar patterns in enrichment across all DEG sets, the other four factors being NHR-80, NHR-28, FOS-1, and PQM-1. The nuclear hormone receptor NHR-80 is robustly expressed in the intestine and regulates fatty acid metabolism via transcriptional control of multiple delta-9 desaturase enzymes (Brock et al., 2006). NHR-28 is highly conserved among nematodes and is expressed in the intestine, pharynx, and hypodermis (Miyabayashi et al., 1999; Reece-Hoyes et al., 2007). The *C. elegans* ortholog of the human bZip FOS transcription factor, FOS-1, is expressed in most tissue types including the hypodermis, intestine, pharynx, and somatic gonad, and is required for proper vulval and uterine development (Hiatt et al., 2009; Oommen and Newman, 2007; Seydoux et al., 1993). The GATA zinc finger transcription factor PQM-1 (paraquat/methylviologen responsive) is primarily expressed in the intestine and is involved in maintenance of proteostasis in response to chronic cellular stress (O'Brien et al., 2018; Shpigel et al., 2019; Tawe et al., 1998). The only previously reported interaction among these 5 factors is between PQM-1 and ELT-2 (MacNeil et al., 2015), and both have also been implicated in regulation of innate immunity responses (Rajan et al., 2019; Shapira et al., 2006). However, as our DEG-TFBS analysis suggests, any or all of these factors may significantly contribute to the transcriptional changes following ectopic *elt-7* expression and are potential regulators of transorganogenesis.

Discussion

The wide-spread ectopic expression of a single transcription factor in *C. elegans* larvae, ELT-7, is capable of directly reprogramming pharyngeal tissue into intestine and the entire uterus undergoes transorganogenesis to become a well-structured intestine-like organ (Riddle et al., 2013; Riddle et al., 2016). Creating an ectopic organ directly from a pre-existing organ is an unprecedented phenomenon and the molecular mechanisms involved in *elt-7*-directed transorganogenesis are unknown. It is also unclear why the somatic gonad and pharynx are particularly susceptible to reprogramming, while other tissues in which *elt-7* is similarly ectopically expressed do not appear to be reprogrammed. Identifying the characteristics of the cellular contexts in which *in vivo* organ reprogramming can be achieved will aid in the discovery of additional transorganogenesis events.

We analyzed the temporal transcriptional profile of *C. elegans* undergoing transorganogenesis and found thousands of genes that are differentially expressed. Enriched among the significantly upregulated DEGs are innate immunity-related genes and more specifically IPR genes (Fig 7). We found that activation of the IPR is restricted to intestine, pharynx, and somatic gonad tissues. Previous reports on the spatial activation of the IPR have only described expression within the intestine, possibly because the *pals-5p::gfp* signal is strongest in intestine cells and more subtle in the other cell types and may have been missed (Bakowski et al., 2014; Botts et al., 2016; Luallen et al., 2015). Transdifferentiation of the pharynx and transorganogenesis of the uterus into intestine correlates both spatially and temporally with IPR gene expression. Therefore, we propose that conversion of cell fate in this context is partially dependent on the predefined ability of a tissue to initiate the IPR (Fig 7A).

Activation of innate immunity pathways have been implicated in other reprogramming events including fibroblast to endothelial cell transdifferentiation driven by exogenous factors (Sayed et al., 2015), fibroblast to neural stem cell reprogramming using a pharmacological cocktail (Zhang et al., 2016), mycobacterium infection-driven reprogramming of Schwann cells (Masaki et al., 2014), and the generation of induced pluripotent stem cells (iPSCs) using Yamanaka factors (Cooke et al., 2014; Lee et al., 2012). It has been proposed that activation of chromatin remodeling factors as part of the cellular inflammatory response may allow for a more open chromatin state that is permissible to binding by ectopic reprogramming transcription factors. This effect has been described as "transflammation" (Cooke et al., 2014; Lee et al., 2012; Shu et al., 2018). However, we also found that other ectopically expressed transcription factors can activate the IPR in the same tissues but do not induce reprogramming, indicating that the IPR alone does not induce a universally plastic state, but rather is specific for transitioning to an intestinal cell fate.

In cells that ectopically express *hsp::elt-7* and the *elt-2p::gfp* intestine reporter transgene, yet which do not show reprogramming phenotypes, we found that genes specific to certain tissues show relatively stable expression throughout the time span we profiled, while genes specific to other tissues, including muscle and hypodermis, show transiently decreased expression (Fig 7B). This suggests that in particular cellular contexts, tissue-specific gene regulation is significantly impacted by ELT-7 activity, but initial cell fate is ultimately maintained, and transcriptional perturbations are eventually corrected for. Cellular processes that maintain stable cell fates are often barriers to cellular reprogramming and can function as *cis*-acting mechanisms, such as epigenetic chromatin modifications that prevent ectopic gene expression (Becker et al., 2016; Cheloufi and Hochedlinger, 2017), and *trans*-acting

mechanisms, such as signaling pathways and ubiquitination, and multiple barriers can have additive effects (Qin et al., 2014; Vierbuchen and Wernig, 2012). Our findings suggest that different modes of cell fate maintenance are in effect in various non-reprogrammed cell types; tissues that are minimally impacted by ectopic ELT-7 activity may primarily rely on *cis*-acting barriers to maintain gene regulatory networks, while tissues that are transiently disrupted by ectopic ELT-7 activity may primarily rely on *trans*-acting barriers to reestablish transcriptional homeostasis.

Previous observation of the intestine-specific reporter transgenes *elt-2p::gfp* and *ifb-2::cfp* following *hsp::elt-7* induction indicated that reprogramming of the pharynx and somatic gonad follows a developmental progression similar to normal embryonic development of the intestine, with activation of *elt-2* occurring first followed by *ifb-2* activation (Riddle et al., 2016). Our gene expression data further supports this conclusion and additionally show that intestine-specifying factors *end-1* and *end-3*, which act upstream of *elt-7* and *elt-2* in embryogenesis, are reactivated during early stages of reprogramming. This redeployment of embryonic factors suggests that transorganogenesis involves recapitulation of normal developmental events. Furthermore, we saw a reciprocal deactivation of the developmental factors *ceh-22* and *fkh-6*, which are specific to the pharynx and somatic gonad respectively.

Upregulated DEGs are significantly correlated with the TFBS profiles of numerous transcription factors. Grouping by similar effect sizes and temporal patterns defines a unique cluster of genes that includes intestine-specific and stress-responsive transcription factors. Some of these factors could facilitate transorganogenesis and may represent a novel transcription factor interaction network. Unexpectedly, we found that in addition to strong

association with upregulated DEGs, the binding site profile for ELT-2 also significantly correlates with downregulated DEGs at 6 and 12 hrs PHS (Fig 6), and there is also significant correlation between downregulated DEGs and intestine-specific genes (Fig 3). It is possible that ELT-2-regulated genes may be downregulated in the endogenous intestine because of a negative feedback mechanism compensating for overexpression of ELT-7 to maintain nominal transcriptional levels. ELT-2 has primarily been described as an activator of transcription (McGhee et al., 2009; Wiesenfahrt et al., 2016) and if it possesses a previously uncharacterized repressive ability, or association with a repressive factor, it could partially explain why it is capable of redirecting cell fates so potently.

The design and analysis of RNA-seq studies are prone to minor experimental differences that can confound the interpretation of differential gene expression (Auer and Doerge, 2010; Fang and Cui, 2011; Lei et al., 2015; Liu et al., 2014; Peixoto et al., 2015). The expression of GFP alone can cause changes in ~1% of genes (Thomas et al., 2005). We used animals with similar, but not identical, transgenic constructs randomly inserted into the genome. Our principal component analysis, however, indicates that our experimental conditions explain the majority of the variation in gene expression. The 0 hrs PHS controls are closely matched for both strains, although there are some initial transcriptomic differences. However, providing the DESeq2 algorithm with our experimental design allows for initial differences to be accounted for (Love et al., 2014). Additionally, we used relatively stringent cutoffs, and our conclusions are drawn primarily from large magnitude effects and highly statistically significant correlations.

One caveat to our temporal transcriptional profiling study is the increasing phenotypic differences between control and experimental strains at later time points. Owing to the

presence of fertilized embryos within adult hermaphrodites only in the control strain at 30 hrs PHS, enrichment of GO terms at late time points need to be carefully interpreted, particularly for processes related to gamete maturation and early embryogenesis. Also, our analysis of correlations between DEGs and TFBS profiles uses a simplistic model for associating transcription factors with target genes and does not attempt to account for variable numbers of binding sites per gene or transcription factor binding outside of the 1.25 kbp promoter regions. Despite these limitations, we saw strong positive correlation with ELT-2, which aligns with our hypotheses for uterus-to-intestine transorganogenesis.

Future work will further examine the relationship between the IPR and ELT-7 driven reprogramming and will aim to determine whether IPR activation is necessary for transorganogenesis. If so, it will be interesting to test whether the IPR can be induced in additional cell types, thereby allowing them to be reprogrammed into intestine. This is currently a challenge, however, as the mechanisms that establish a tissue's ability to activate the IPR and the primary signals for activation of IPR genes are unknown. Based on our earlier work, it is plausible that PHA-4 activity during development may prime intestine, pharynx, and somatic gonad tissues for IPR activation (Riddle et al., 2016), and our current findings suggest that PQM-1, a known stress response factor (O'Brien et al., 2018; Rajan et al., 2019; Shpigel et al., 2019), could be playing a role in IPR gene activation post-embryonically following ectopic *elt-7* expression.

Our study further illustrates the changes in cell fate of the pharynx and somatic gonad into intestine, yet it is still to be determined whether these tissues have undergone complete transdifferentiation or if there are significant remnants of their initial identities by 30 hours after *elt-7* expression. Incomplete reprogramming is a major hurdle to the clinical application

of cell therapies (Ebrahimi, 2017; Taguchi and Yamada, 2017; Srivastava and DeWitt, 2016). This could be resolved using a transcriptional profiling method with greater spatial resolution, either by dissecting uterus tissue for RNA extraction or utilizing single cell RNA sequencing (Wagner et al., 2016), for example.

It is still unclear as to why non-reprogrammed tissues vary significantly in their transcriptional responses to ectopic ELT-7 activity. We speculate that differences in chromatin structure and similarity between the overall nuclear topology of intestine cells and target cell types may account for the ability of ELT-7 to modulate transcription in certain tissues. This hypothesis could be tested with *C. elegans* atlases of tissue-type chromatin accessibility, such as has been achieved with fruit fly and mouse embryos (Cusanovich et al., 2018a; Cusanovich et al., 2018b), and chromosomal conformation (Ramani et al., 2017). Future cross comparisons of our temporal transcriptional profiling data with the inevitable emergence of such resources will likely reveal additional insight into the mechanisms underlying transorganogenesis.

In summary, our temporal transcriptional profiling of *C. elegans* undergoing transorganogenesis represents the first gene expression map of whole organ reprogramming *in vivo*. Our results indicate that in order to effectively reprogram an organ, certain commonalities may need to exist between the starting and ending tissue types. The shared ability to activate an innate immune response to intracellular pathogen infection implicates transflammation as a potential critical cellular process in *elt-7*-directed reprogramming. The transient repression of tissue-specific gene expression in non-reprogrammed tissue types reveals that even if *elt-7* expression causes no obvious cellular phenotypes, there could be significant disruption of the transcriptional state. This may be analogous to off-target effects

of gene therapy and emphasizes the need for detailed molecular characterization of tissues for safe application of *in vivo* reprogramming. These data will enable further targeted research into the mechanisms of transorganogenesis and may lead to the discovery of similar phenomena in *C. elegans* and other organisms.

Materials and Methods

Strains used and culturing.

C. elegans strains were grown at 20^{o}C on NGM media seeded with OP50 bacteria as previously described (Brenner, 1974; Stiernagle, 2006). Experiments were performed at 20^{o}C unless specified otherwise. JR3642 wIs125[hsp::elt-7]; rrIs1[elt-2p::gfp]; kcIs6[ifb-2::cfp], CL2070 dvIs70[hsp-16.2p::gfp + rol-6(su1006)], ERT54 jyIs8[pals-5p::gfp + myo-2p::mCherry] X., JR4031 kcIs6[ifb-2::cfp]; jyIs8[pals-5p::gfp + myo-2p::mCherry] X; wIs125[hs::elt-7 + pRF4 (rol-6(su1004))], JR4033 jyIs8[pals-5p::gfp + myo-2p::mCherry] X; caIs6[hs::elt-1 + Rol], JR4034 jyIs8[pals-5p::gfp + myo-2p::mCherry] X; [hs::hlh-1 + Rol], JR4038 jyIs8[pals-5p::gfp + myo-2p::mCherry] X; kcIs6[ifb-2::cfp]; wIs76[hs::end-3 + Rol], JR4040 jyIs8[pals-5p::gfp + myo-2p::mCherry] X; kcIs6[ifb-2::cfp]; wIs47[hs::end-1 + Rol].

Heat shock, RNA purification, and sequencing library prep.

Heat shock was performed by placing synchronized, early L4 stage worms growing on NGM plates in a 33^{0}C incubator for 30 minutes. For each sample approximately 5,000-20,000 worms were harvested by washing off plates into M9 buffer, washed twice, rotated for 15 minutes at room temperature to allow most bacteria to be excreted by the worms, then washed two more times. Worms were resuspended in TRI Reagent (Zymo Research R2050-

1) vortexed for 2 minutes, and then flash frozen in liquid nitrogen. Worm lysis was achieved by rapidly freeze-thawing samples 3-4 times, alternating submersion in a 37^0C DEPC-treated water bath and liquid nitrogen. Worm debris was pelleted and nucleotide-containing supernatant was transferred to a Zymogen column. Total RNA was purified using a Zymo Research Direct-zol RNA Mini Prep kit (cat no: R2050) according to manufacturer's protocol. High quality RNA was obtained as determined by Agilent TapeStation RIN scores.

Sequencing libraries were built using an Illumina TruSeq Stranded mRNA sample prep kit following the LS protocol using 1-4 µg of total RNA input.

Libraries were sequenced by UCSB Genetics Core staff on an Illumina NextSeq 500 instrument. Three hours post heat shock samples were sequenced on a mid-output 300 cycle paired-end run, all other samples were sequenced across two runs using high-output 75 cycle single-end reads.

Sequence read QC, alignments, quantification, and differential expression analysis

Data processing was performed in Galaxy (usegalaxy.org) using a custom workflow (Afgan et al., 2018; Blankenberg et al., 2014). Raw fastq files were concatenated for each sample and trimmed using Trimmomatic (Galaxy Version 0.36.3) with the parameters: ILLUMINACLIP:/TruSeq3-SE.fa:2:30:10 SLIDINGWINDOW:4:10 LEADING:5 TRAILING:5 MINLEN:50 AVGQUAL:20 for single reads and: ILLUMINACLIP:/TruSeq3-PE.fa:2:30:10 MINLEN:100 AVGQUAL:20 SLIDINGWINDOW:4:10 LEADING:5 TRAILING:5 for paired-end reads (Bolger et al., 2014). Pseudoalignment and abundance quantification of RNA-seq reads was performed using Salmon (Galaxy Version 0.8.2) to the *C. elegans* WS263 reference transcriptome with default parameters (Patro et al., 2017). Differential gene expression analysis was peformed using DESeq2 v1.20.0 following the

time-series variation to the standard workflow (Love et al., 2014). Principal component analysis was performed on regularized-logarithm transformed count data.

Spatial gene expression analysis

Sets of genes previously determined to have selectively enriched expression (> 2x, FDR < 5%) in a particular cell-type versus a reference sample of all cells at the same stage were obtained from https://www.vanderbilt.edu/wormdoc/wormmap/Selectively_enriched_genes.html (Spencer et al., 2011). Statistical analysis of correlation between up- and downregulated differentially expressed genes and cell-type enriched genes was performed by applying Fisher's exact test on 2x2 contingency tables comparing the number of genes that are both DE and cell-type expressed, DE but not cell-type expressed, not DE and cell-type expressed, and not DE and not cell-type expressed within the complete set of 20,094 genes quantified.

Gene ontology enrichment analysis

Venn diagrams and clusters for up- and downregulated DEGs were made using Venny 2.1 (Oliveros, J.C. 2007-2015) https://bioinfogp.cnb.csic.es/tools/venny/index.html. Gene ontology enrichment analysis was performed using the clusterProfiler v3.8.0 Bioconductor package implemented in R v3.5.0 (Yu et al., 2012). Dotplots shown in Fig 4 were filtered for semantic similarity and only the top terms for each cluster are shown.

Time-lapse imaging of intracellular pathogen response

Worms were mounted on 4% agar pads on microscope slides and immobilized with ~5 µM levamisole. Worms were heat shocked by placing slides in a 33^0C incubator for 30 minutes and then mounted on a microscope for time-lapse imaging. Images were captured

using a Nikon Eclipse Ti microscope equipped with a Hamamatsu flash Orca 2.8 camera and a 20X objective lens. Image processing, including resizing, cropping, LUT adjustment, and channel merging, was done using Nikon Elements, Fiji (Image J), and Inkscape software.

Transcription factor ChIP-seq and DEG analysis

Genome-wide DNA binding site profiles for 217 GFP-tagged *C. elegans* transcription factors was obtained from the model organism encyclopedia of regulatory networks (modERN) study (Kudron et al., 2017). Compiled ChIP-seq data contained 667,924 peaks, which were mapped to putative promoter regions using a custom R script, resulting in 444,136 transcription factor binding site-promoter associations. Promoter regions were defined as -1,000 to +250 bp relative to the transcription start sites for the 20,094 protein coding genes analyzed in this study.

Statistical analysis of correlations between genes bound by a given transcription factor, and up- or downregulated DEGs at 3, 6, 12, and 30 hrs PHS was performed using Fisher's exact test.

Fig 1.

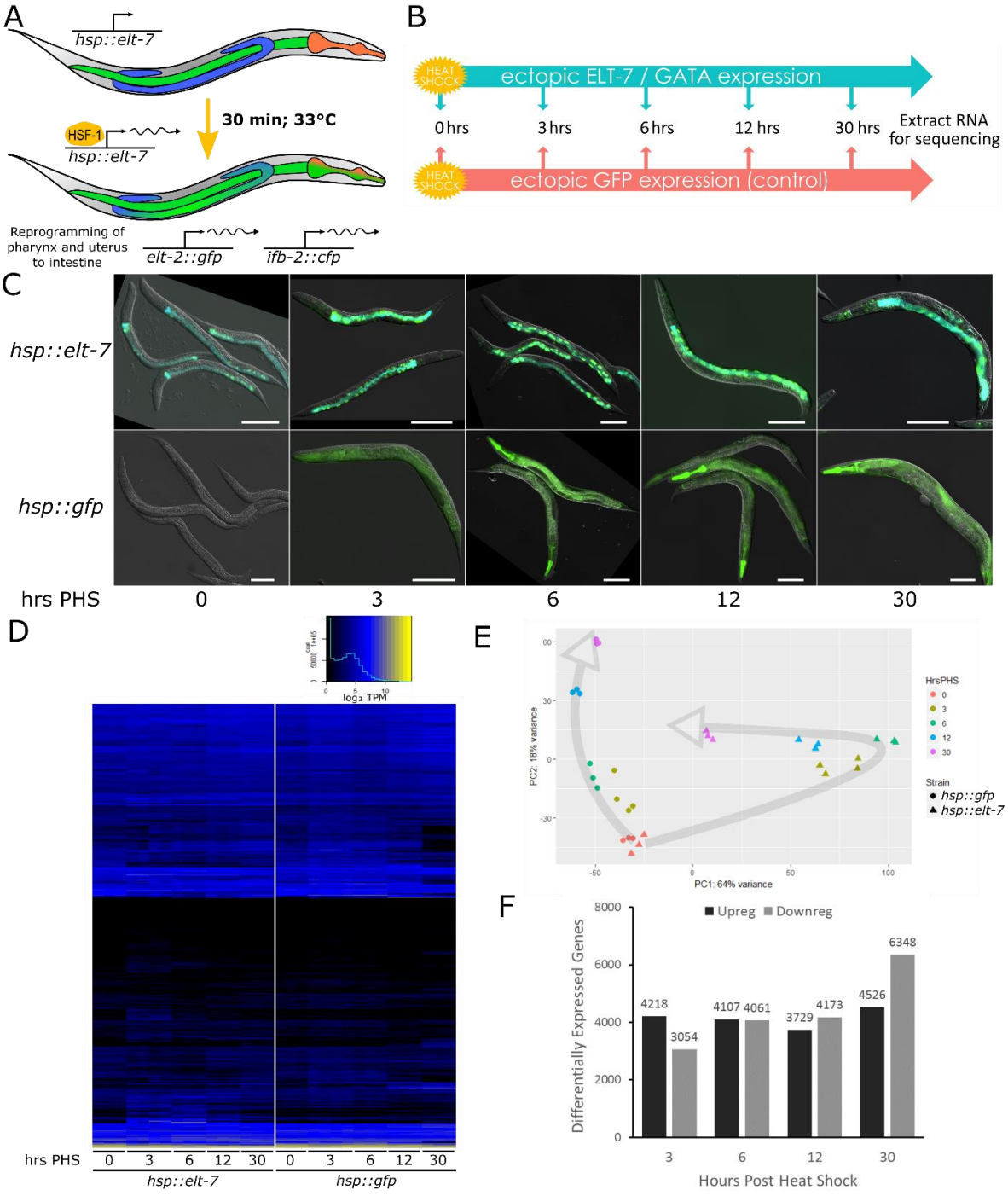

A
hsp::elt-7
30 min; 33°C
HSF-1
hsp::elt-7
Reprogramming of pharynx and uterus to intestine
elt-2::gfp
ifb-2::cfp
B
HEAT SHOCK
ectopic ELT-7 / GATA expression
0 hrs
3 hrs
6 hrs
12 hrs
30 hrs
Extract RNA for sequencing
HEAT SHOCK
ectopic GFP expression (control)
C
hsp::elt-7
hsp::gfp
hrs PHS
0
3
6
12
30
D
log2 TPM
hrs PHS
0 3 6 12 30 0 3 6 12 30
hsp::elt-7
hsp::gfp
E
PC2: 18% variance
PC1: 64% variance
HrsPHS
0
3
6
12
30
Strain
hsp::gfp
hsp::elt-7
F
Upreg
Downreg
Differentially Expressed Genes
8000
6000
4000
2000
0
4218
3054
4107
4061
3729
4173
4526
6348
3
6
12
30
Hours Post Heat Shock

Fig 1. Temporal transcriptional profiling and differential gene expression of *C. elegans* undergoing ectopic *elt-7* directed transdifferentiation and transorganogenesis.

(A) A *C. elegans* strain with a transgene containing the endoderm-specific GATA-type transcription factor *elt-7* under the control of a heat shock response promoter is ubiquitously expressed in all cells after a brief temperature shift. Within the cellular context of the differentiated pharynx and the developing uterus, *elt-7* induces cellular reprogramming to an intestine cell fate. The reprogrammed pharynx and uterus stably express the intestine tissue specific markers *elt-2::gfp* and *ifb-2::cfp*. Other tissues transiently express *elt-2::gfp* after heat shock but do not become reprogrammed. (B) Total RNA was extracted for high throughput mRNA-sequencing from synchronous cultures of L4 stage worms at multiple time points throughout the *in vivo* cellular reprogramming process, as well as from a control strain that expresses a *hsp::gfp* transgene, with 3-4 biological replicates for each condition. (C) Representative fluorescence and DIC microscopy images of worms from each sample used for temporal transcriptional profiling. Immediately before heat shock (0 hours post heat shock (hrs PHS)) *elt-2::gfp* (green) and *ifb-2::cfp* (blue) are expressed only in the intestine. Beginning approximately 3 hrs PHS *elt-2::gfp* is expressed throughout the worm and becomes increasingly restricted to the pharynx and uterus at later time points. Ectopic *ifb-2::cfp* first becomes visible approximately 6 hrs PHS, and by 30 hrs PHS is clearly visible in the uterus, which has been reprogrammed into a secondary intestine-like organ. In the *hsp::gfp* control strain, no GFP is visible at 0 hrs PHS. At 3 hrs PHS GFP initially becomes visible throughout all worms and remains visible at all later time points. Brightness and contrast have been adjusted to visualize spatial fluorescence patterns. Scale bar: 100 μm. (D) mRNA-seq was performed across three runs on an Illumina NextSeq500. Pseudoalignment to

the WS263 reference transcriptome and quantification of gene expression for 20,094 genes was performed using Salmon (version 0.8.2). Heatmap displays the normalized $\log_2$ transformed transcripts per million (TPM) for all 32 mRNA-seq samples. (E) Principle component analysis highlights close clustering of pre-heat shock samples, sample replicates, and separation between strains into distinct expression paths (arrows). The *hsp::gfp* control strain samples are ordered along PC2 by time point, and therefore the y-axis correlates with developmental gene expression changes. The *hsp::elt-7* samples are separated by PC1 along a curved path with greatest separation from the control at 6 hrs PHS, indicating a dynamic transcriptional response to ectopic ELT-7. (F) Differential gene expression analysis of *hsp::elt-7* versus *hsp::gfp* was performed using DESeq2. On average, over 4,000 genes were detected as significantly upregulated and downregulated at each time point (FDR < 0.01).

Table 1. Summary of reads obtained for mRNA-seq sample libraries.

			hsp::elt-7		*hsp::gfp*	
Hrs PHS	Replicate	SE/PE	Total Reads	Trimmed Reads	Total Reads	Trimmed Reads
0	1	SE	61,245,563	61,170,600	22,848,404	22,785,563
0	2	SE	71,111,147	70,992,518	18,815,620	18,761,475
0	3	SE	21,326,572	21,282,606	16,590,803	16,560,047
3	1	PE	25,172,963	20,184,434	25,120,129	20,498,671
3	2	PE	23,223,890	18,826,550	22,299,853	17,662,326
3	3	PE	22,413,501	17,872,167	23,261,219	18,867,519
3	4	PE	22,365,672	17,590,590	22,218,267	18,236,942
6	1	SE	20,562,929	20,532,185	49,429,326	49,387,239
6	2	SE	59,057,122	59,014,361	61,702,785	61,560,194
6	3	SE	66,391,376	66,334,135	64,818,789	64,744,015
12	1	SE	37,812,327	37,769,576	38,831,700	38,805,605
12	2	SE	42,587,999	42,530,347	73,755,973	73,692,370
12	3	SE	41,852,940	41,801,924	56,013,566	55,974,981
30	1	SE	72,377,148	72,337,022	49,877,266	49,852,333
30	2	SE	54,338,024	54,311,567	69,797,347	69,743,673
30	3	SE	58,544,598	58,499,528	71,434,388	71,403,250

Hrs PHS, hours post heat shock. SE, single end reads. PE, paired end reads.

Fig 2.

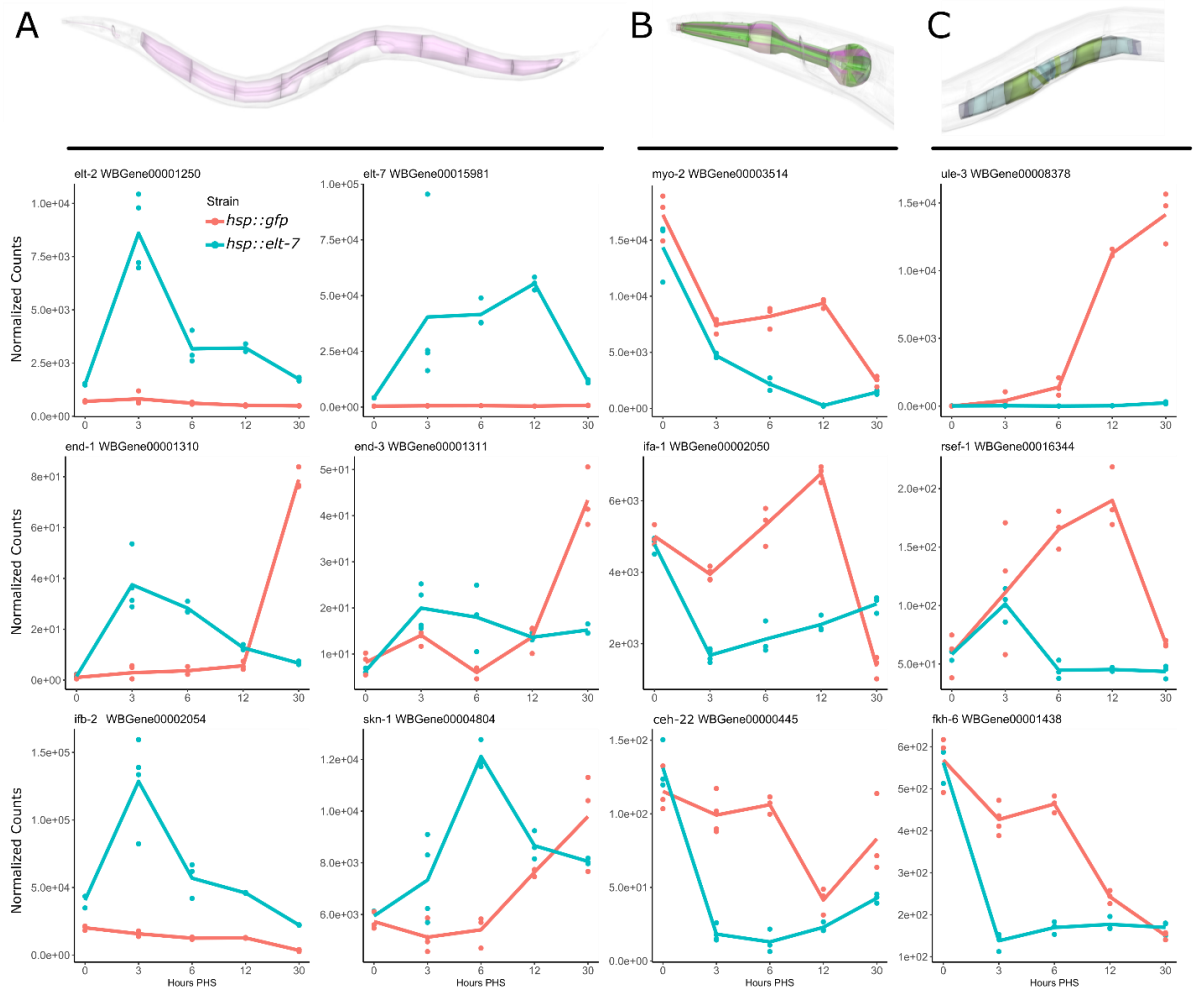

A
B
C
elt-2 WBGene00001250
elt-7 WBGene00015981
myo-2 WBGene00003514
ule-3 WBGene00008378
end-1 WBGene00001310
end-3 WBGene00001311
ifa-1 WBGene00002050
rsef-1 WBGene00016344
ifb-2 WBGene00002054
skn-1 WBGene00004804
ceh-22 WBGene00000445
fkh-6 WBGene00001438
Strain
hsp::gfp
hsp::elt-7
Normalized Counts
Hours PHS

Fig 2. Cellular reprogramming-related gene expression changes are captured by transcriptional profiling and indicate redirection of developmental factors.

(A) Genes known to be specifically involved in endoderm development are upregulated during transorganogenesis. *elt-2* and *elt-7* are intestine-differentiating transcription factors expressed in the E-cell lineage from mid embryogenesis throughout adulthood. *end-1* and *end-3* are redundant transcription factors that specify the endoderm. They are normally expressed only transiently during early embryogenesis but are reactivated after ectopic *elt-7* expression. *ifb-2* is an intestine-specific intermediate filament protein which localizes to the terminal web of the intestinal lumen. *skn-1* is a critical transcription factor in establishing the E-cell lineage as well as functioning post-embryonically as a stress response factor. (B) Transdifferentiation of the pharynx is evidenced by a rapid downregulation of pharyngeal genes. The pharynx expressed structural genes *myo-2* and *ifa-1*, and the pharynx-specific homeobox transcription factor *ceh-22* are strongly downregulated after ectopic *elt-7* expression. (C) Uterus-to-intestine transorganogenesis results in downregulation of developmental factors and failure to activate uterus maturation factors. *ule-3* (Uterine Lumen Expressed) is one of multiple factors strongly expressed in the adult uterus that have minimal expression detected after ectopic *elt-7* expression. *rsef-1* (RaS and EF hand domain containing homolog) is normally expressed in the uterine toroids and functions in proper development of the uterine seam cell. The forkhead box transcription factor *fkh-6* is involved in gonadogenesis and is rapidly downregulated following ectopic *elt-7* expression. Anatomic expression pattern images courtesy of wormbase.org.

Fig 3.

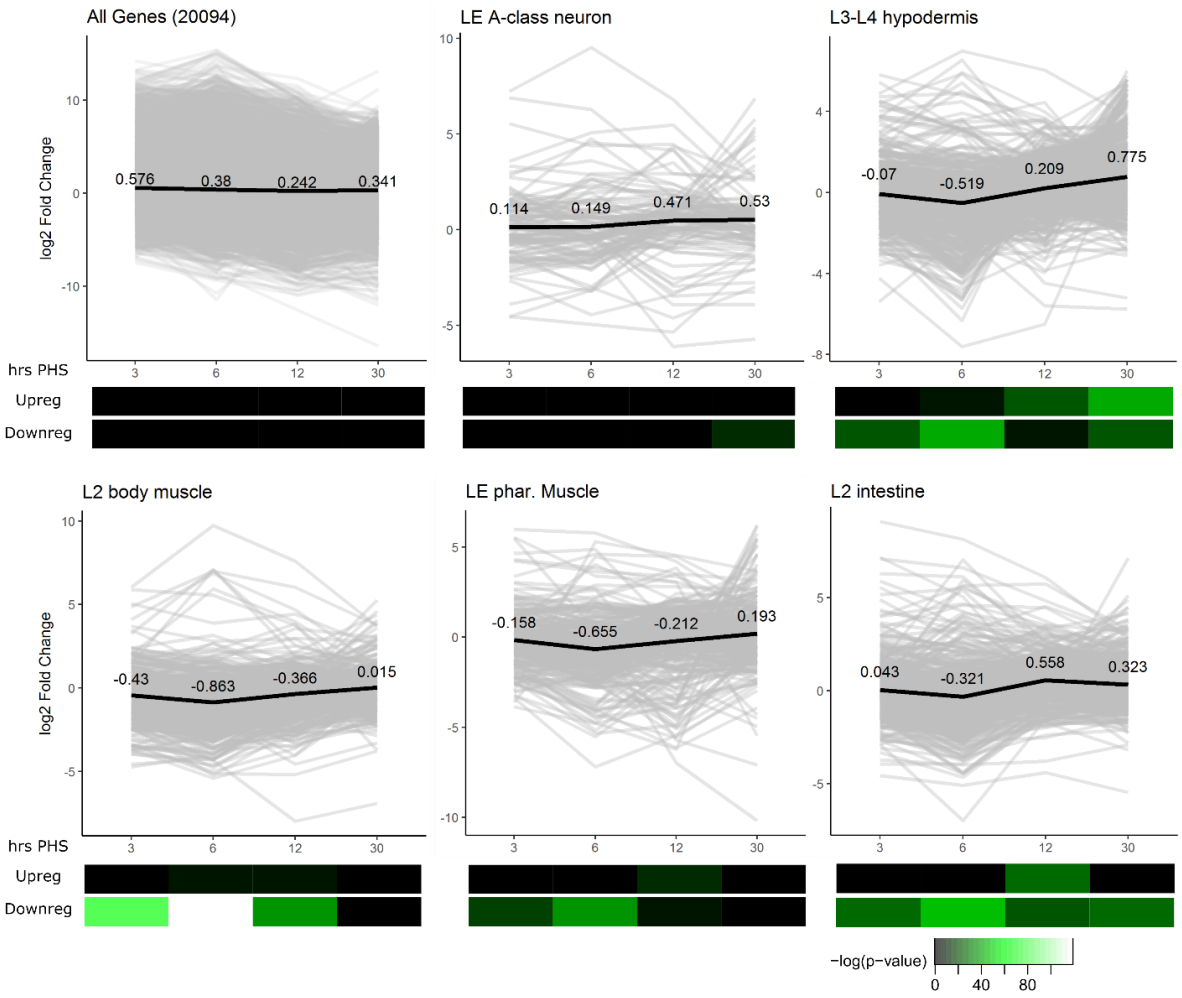

All Genes (20094)
LE A-class neuron
L3-L4 hypodermis
L2 body muscle
LE phar. Muscle
L2 intestine
log2 Fold Change
0.576
0.38
0.242
0.341
0.114
0.149
0.471
0.53
-0.07
-0.519
0.209
0.775
-0.43
-0.863
-0.366
0.015
-0.158
-0.655
-0.212
0.193
0.043
-0.321
0.558
0.323
hrs PHS
Upreg
Downreg
−log(p−value)
0
40
80

Fig 3. Tissue specific analysis of gene expression dynamics reveals transient transcriptional perturbations in non-transdifferentiated cell types.

Sets of genes found to be enriched for expression in specific tissue types by a previous *C. elegans* spatial transcriptional profiling study were plotted by $\log_2$ fold change (*hsp::elt-7* vs *hsp::gfp* gene expression) at 3, 6, 12, and 30 hrs PHS. Each gray line represents an individual gene, and black lines show the mean differential expression path with corresponding fold change values. Separate statistical testing was performed to calculate the likelihood that a significant proportion of genes within an enriched tissue set are also present in either the upregulated (Upreg) or downregulated (Downreg) set of differentially expressed genes (DEGs) for each time point using Fisher's Exact Test. Heatmaps display the $-\log_{10}$ p-values for each gene set comparison. A complete fold change plot of all 20,094 genes shows that despite a range of over 2,000-fold differential gene expression changes, overall gene expression is only mildly upregulated (<1.5 fold) across all time points. A-class neurons exemplify a tissue type with large fold changes for some genes but average changes comparable to base line and only weak statistical correlation between Downreg DEGs at 30 hrs PHS. *hsp::elt-7* is expressed ectopically in all tissues including skin and muscle, and although they do not transdifferentiate into intestine, hypodermal and body muscle tissue enriched gene sets have transiently decreased overall expression with the greatest fold change decrease at 6 hrs PHS and the greatest fold change increase at 30 hrs PHS. Pharynx muscle enriched genes also display a transient average fold change decrease that may be attributable to partial or incomplete reprogramming of pharynx muscle tissues. Intestine enriched genes are downregulated on average at 6 hrs PHS and upregulated at 12 and 30 hrs PHS and may

indicate distinct transcriptional phases of the cellular reprogramming process. LE, late embryo; L2, L3, L4, larval stage 2, 3 ,4.

Fig 4.

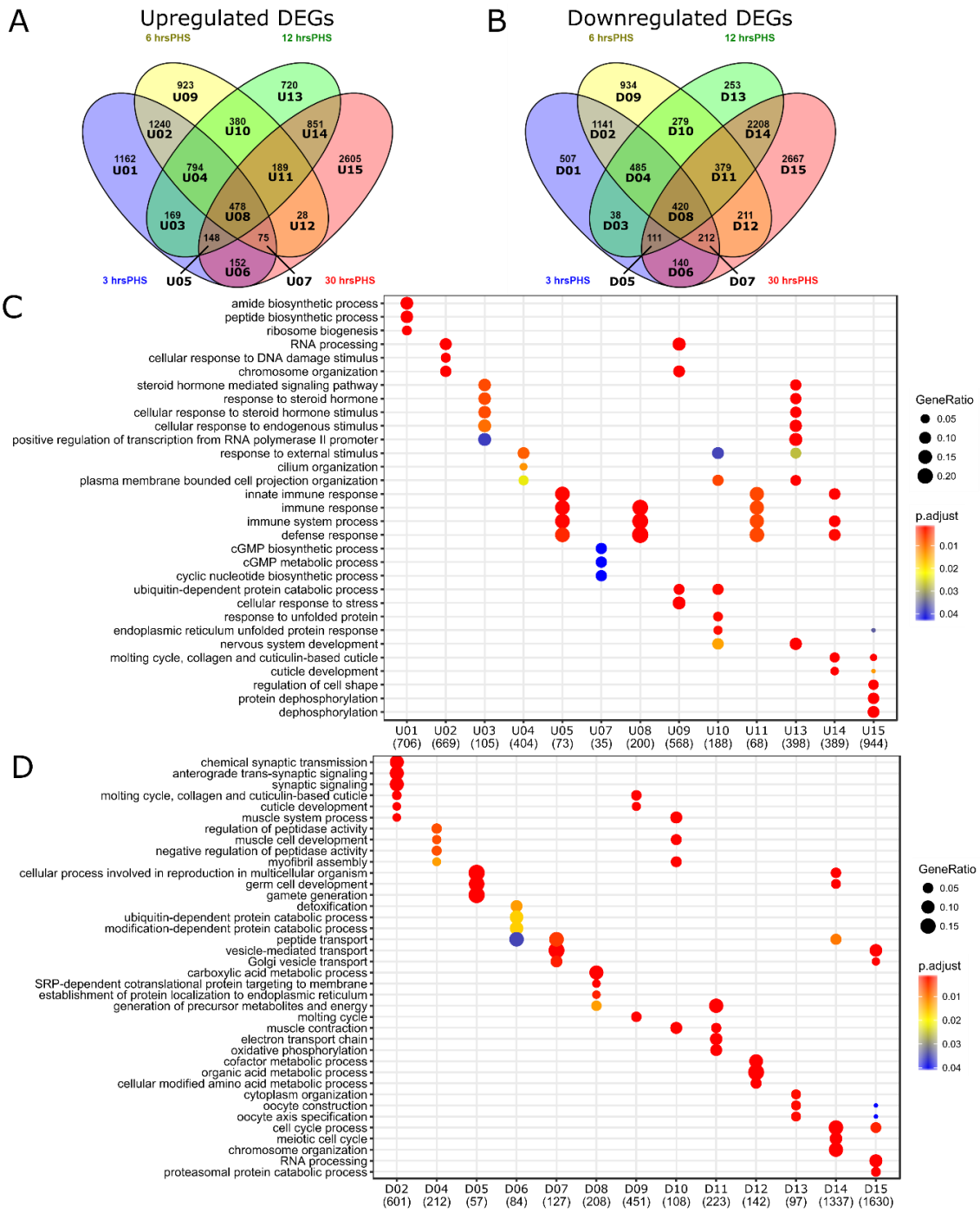

A
Upregulated DEGs
6 hrsPHS
12 hrsPHS
923
U09
720
U13
1240
U02
380
U10
851
U14
1162
U01
794
U04
189
U11
2605
U15
169
U03
478
U08
28
U12
148
75
152
U06
3 hrsPHS
U05
U07
30 hrsPHS
B
Downregulated DEGs
6 hrsPHS
12 hrsPHS
934
D09
253
D13
1141
D02
279
D10
2208
D14
507
D01
485
D04
379
D11
2667
D15
38
D03
420
D08
211
D12
111
212
140
D06
3 hrsPHS
D05
D07
30 hrsPHS
C
amide biosynthetic process
peptide biosynthetic process
ribosome biogenesis
RNA processing
cellular response to DNA damage stimulus
chromosome organization
steroid hormone mediated signaling pathway
response to steroid hormone
cellular response to steroid hormone stimulus
cellular response to endogenous stimulus
positive regulation of transcription from RNA polymerase II promoter
response to external stimulus
cilium organization
plasma membrane bounded cell projection organization
innate immune response
immune response
immune system process
defense response
cGMP biosynthetic process
cGMP metabolic process
cyclic nucleotide biosynthetic process
ubiquitin-dependent protein catabolic process
cellular response to stress
response to unfolded protein
endoplasmic reticulum unfolded protein response
nervous system development
molting cycle, collagen and cuticulin-based cuticle
cuticle development
regulation of cell shape
protein dephosphorylation
dephosphorylation
U01 (706)
U02 (669)
U03 (105)
U04 (404)
U05 (73)
U07 (35)
U08 (200)
U09 (568)
U10 (188)
U11 (68)
U13 (398)
U14 (389)
U15 (944)
GeneRatio
0.05
0.10
0.15
0.20
p.adjust
0.01
0.02
0.03
0.04
D
chemical synaptic transmission
anterograde trans-synaptic signaling
synaptic signaling
molting cycle, collagen and cuticulin-based cuticle
cuticle development
muscle system process
regulation of peptidase activity
muscle cell development
negative regulation of peptidase activity
myofibril assembly
cellular process involved in reproduction in multicellular organism
germ cell development
gamete generation
detoxification
ubiquitin-dependent protein catabolic process
modification-dependent protein catabolic process
peptide transport
vesicle-mediated transport
Golgi vesicle transport
carboxylic acid metabolic process
SRP-dependent cotranslational protein targeting to membrane
establishment of protein localization to endoplasmic reticulum
generation of precursor metabolites and energy
molting cycle
muscle contraction
electron transport chain
oxidative phosphorylation
cofactor metabolic process
organic acid metabolic process
cellular modified amino acid metabolic process
cytoplasm organization
oocyte construction
oocyte axis specification
cell cycle process
meiotic cell cycle
chromosome organization
RNA processing
proteasomal protein catabolic process
D02 (601)
D04 (212)
D05 (57)
D06 (84)
D07 (127)
D08 (208)
D09 (451)
D10 (108)
D11 (223)
D12 (142)
D13 (97)
D14 (1337)
D15 (1630)
GeneRatio
0.05
0.10
0.15
p.adjust
0.01
0.02
0.03
0.04

Fig 4. Gene ontology (GO) enrichment profiles of DEG clusters identifies the temporal sequence of diverse cellular processes modulated during *in vivo* cellular reprogramming.

(A) Overlaps between upregulated DEGs at 3, 6, 12, and 30 hrs PHS were used to define 15 clusters of non-redundant genes representing unique gene expression paths for GO enrichment analysis (U01-U15). (B) Downregulated DEG clusters (D01-D15) were defined in the same manner. (C) Dotplot of the top Biological Process GO terms enriched for each upregulated DEG cluster. Dot column labels correspond to gene clusters in (A) and the number of genes with sufficient annotation that were used for GO enrichment analysis in each cluster. Dot size represents the proportion of genes in the cluster annotated with each term. Dot color represents the adjusted p-value with warmer shades indicating greater significance of enrichment. Most biological processes are only enriched among one or two DEG clusters with the exception of immunity related terms that are enriched throughout the reprogramming time course. (D) Dotplot of the top Biological Process GO terms enriched for each downregulated DEG cluster. Similarly to upregulated GO terms, there is little redundancy of enrichment between clusters, indicating reproducible temporal specificity of cellular perturbations associated with *elt-7* directed cellular reprogramming. Clusters not shown in dotplots had no significant GO enrichments.

Fig 5.

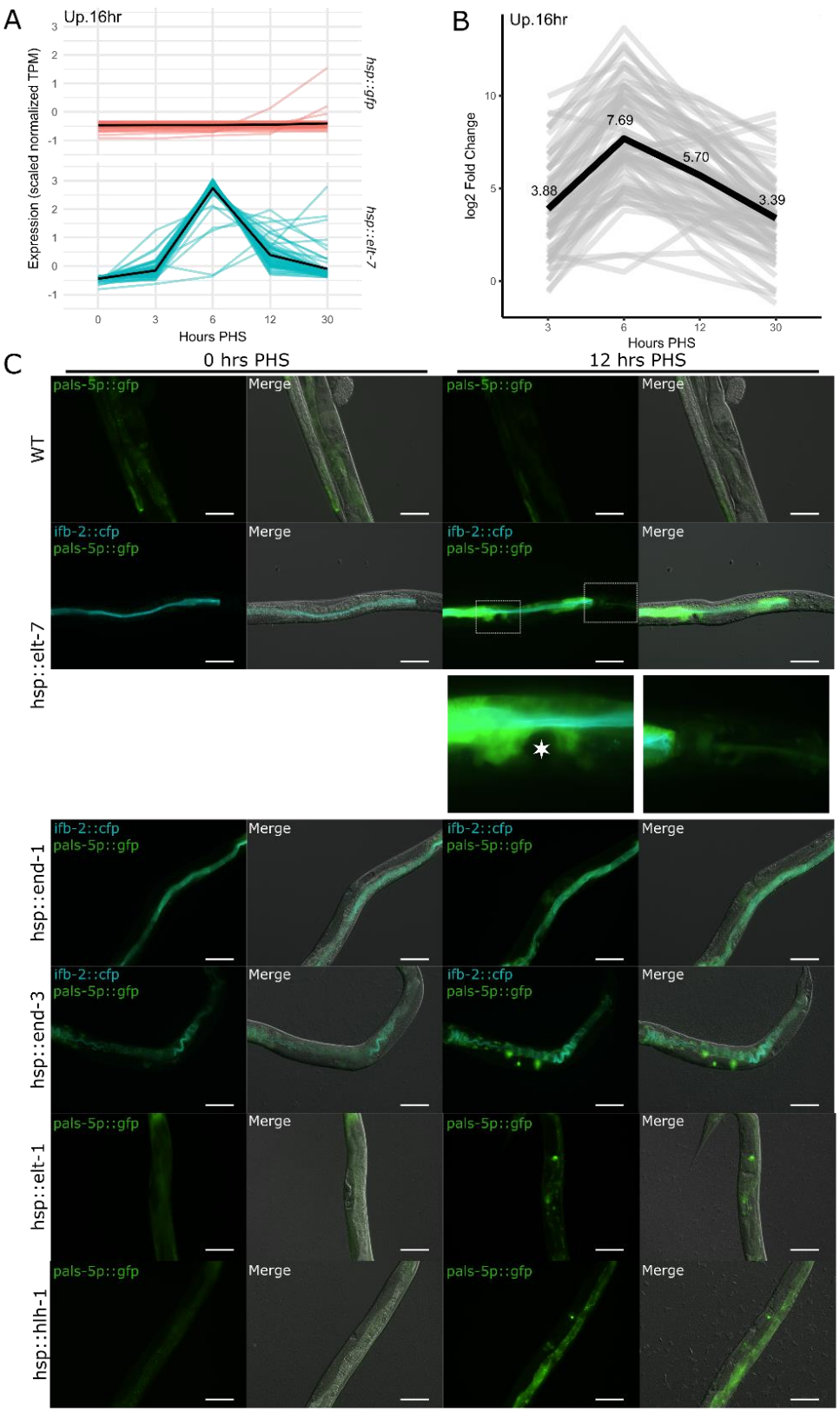

A
Up.16hr
Expression (scaled normalized TPM)
hsp::gfp
hsp::elt-7
Hours PHS
B
Up.16hr
log2 Fold Change
7.69
5.70
3.88
3.39
Hours PHS
C
0 hrs PHS
12 hrs PHS
WT
hsp::elt-7
hsp::end-1
hsp::end-3
hsp::elt-1
hsp::hlh-1
pals-5p::gfp
ifb-2::cfp
Merge

Fig 5. Ectopic *elt-7* expression induces an intracellular pathogen-like transcriptional response within tissues susceptible to reprogramming to intestine fate.

Differentially expressed genes from this study were found to statistically correlate with differential gene expression changes following intracellular pathogen infection by microsporidia and viruses. (A) Temporal gene expression profiles for 68 genes previously identified as significantly upregulated 16 hours post infection by *N. parisii* (Bakowski et al., 2014). Top plot shows very low basal expression for all 68 genes in control samples. Bottom plot shows strong induction in expression of intracellular pathogen response (IPR) genes after ectopic *elt-7* expression with a peak in transcription around 6 hrs PHS. Black line shows mean TPM. (B) *hsp::elt-7* vs *hsp::gfp* $\log_2$ fold changes for 68 genes upregulated 16 hours post infection by *N. parisii*. Black line shows mean fold change. IPR genes are over 200-fold upregulated on average at 6 hrs PHS. (C) Spatial activation of IPR after ectopic transcription factor expression. The IPR reporter construct *pals-5p::gfp* was crossed into various heat shock promoter::transcription factor expressing strains. Time-lapse fluorescence and DIC imaging of individual worms was performed to identify the spatial and temporal pattern of *pals-5p::gfp* expression after heat shock. Representative images from 0 and 12 hrs PHS, when *pals-5p::gfp* expression was strongest, are shown for each strain. The 30 minute, 33^0C heat shock treatment does not induce *pals-5p::gfp* expression in a wild type background. 12 hours post ectopic *elt-7* expression, *pals-5p::gfp* is strongly induced throughout the intestine, and in addition, *pals-5p::gfp* is observed in somatic gonad tissue surrounding the vulva (star) and anterior to the intestine in the pharynx (insets). Ectopic *end-1* produces a weak induction of *pals-5p::gfp* with a similar expression pattern to *elt-7*. Ectopic *end-3, elt-1,* and *hlh-1* induce *pals-5p::gfp* weakly in the intestine and in the somatic gonad, but more focused in

regions corresponding to the spermatheca and vulva. Brightness and contrast have been adjusted to illustrate spatial fluorescence expression patterns. Scale bars, 100 μm.

Fig 6.

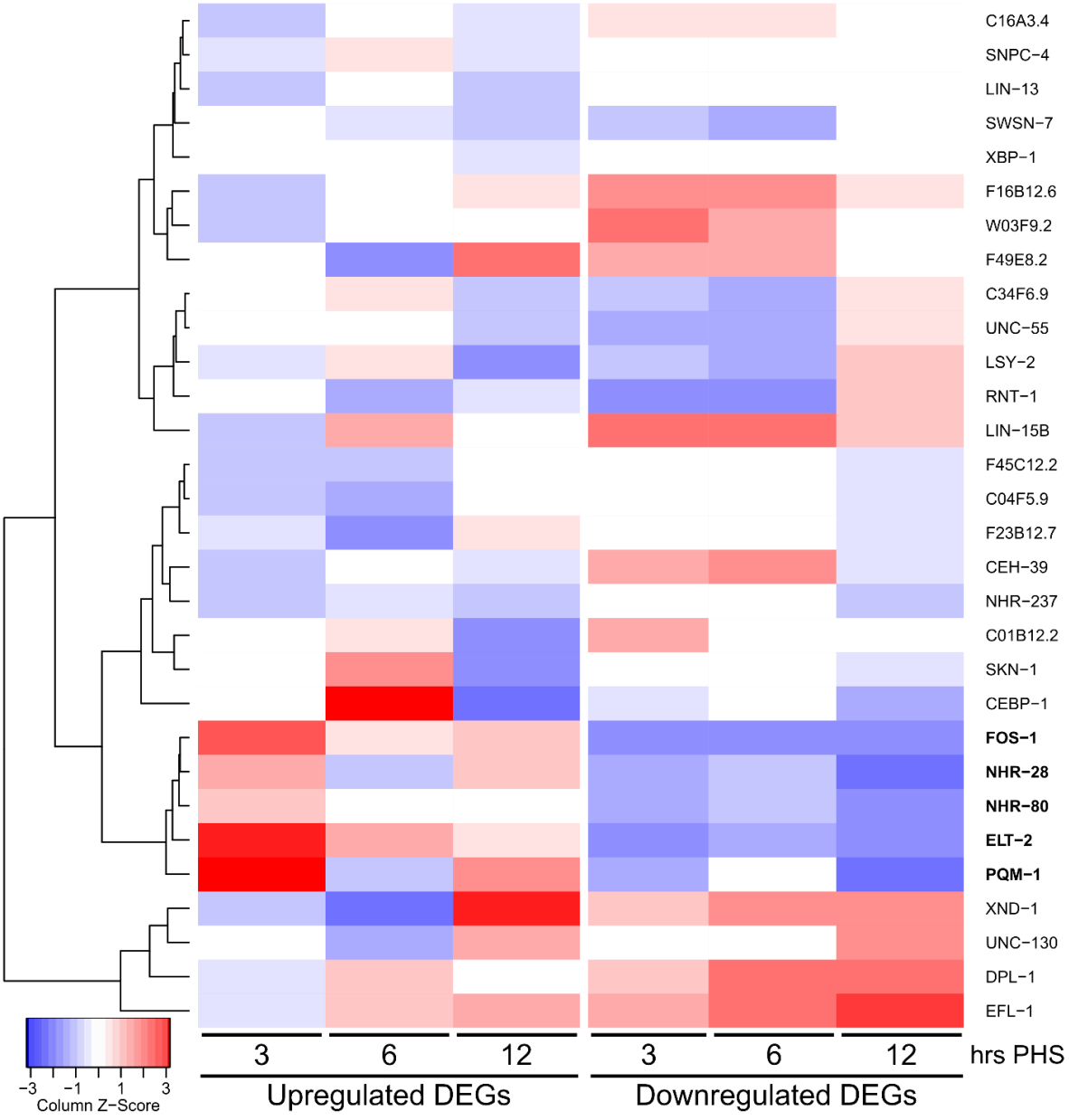

C16A3.4
SNPC-4
LIN-13
SWSN-7
XBP-1
F16B12.6
W03F9.2
F49E8.2
C34F6.9
UNC-55
LSY-2
RNT-1
LIN-15B
F45C12.2
C04F5.9
F23B12.7
CEH-39
NHR-237
C01B12.2
SKN-1
CEBP-1
FOS-1
NHR-28
NHR-80
ELT-2
PQM-1
XND-1
UNC-130
DPL-1
EFL-1
3
6
12
3
6
12
hrs PHS
Upregulated DEGs
Downregulated DEGs
−3
−1
1
3
Column Z-Score

Fig 6. Correlation of transcription factor binding site profiles with *elt-7* induced differentially expressed genes identifies potential regulators of transorganogenesis. Transcription factor chromatin immunoprecipitation followed by deep sequencing (TF ChIP-seq) data for 217 *C. elegans* TFs from the recent modERN study (Kudron et al., 2017) was mapped to putative promoter regions for all 20,094 protein coding genes quantified in this study and cross referenced to up- and downregulated DEG sets for significant enrichment using Fisher's Exact test. Heatmap displays $-\log_{10}$ p-values scaled by column ranges for the top 30 TFs with the most significant correlations to DEGs upregulated at 3, 6, and 12 hrs PHS, with red indicating greatest significance. Dendrogram clustering of TFs is by row similarity. One of the TFs that correlates most significantly with upregulated DEGs is ELT-2, which is the primary activator of intestine differentiation. Four other TFs have similar temporal correlations with DEGs as ELT-2 (bolded) and may be functioning as important transcriptional activators during pharynx-to-intestine transdifferentiation and/or uterus-to-intestine transorganogenesis.

Fig 7.

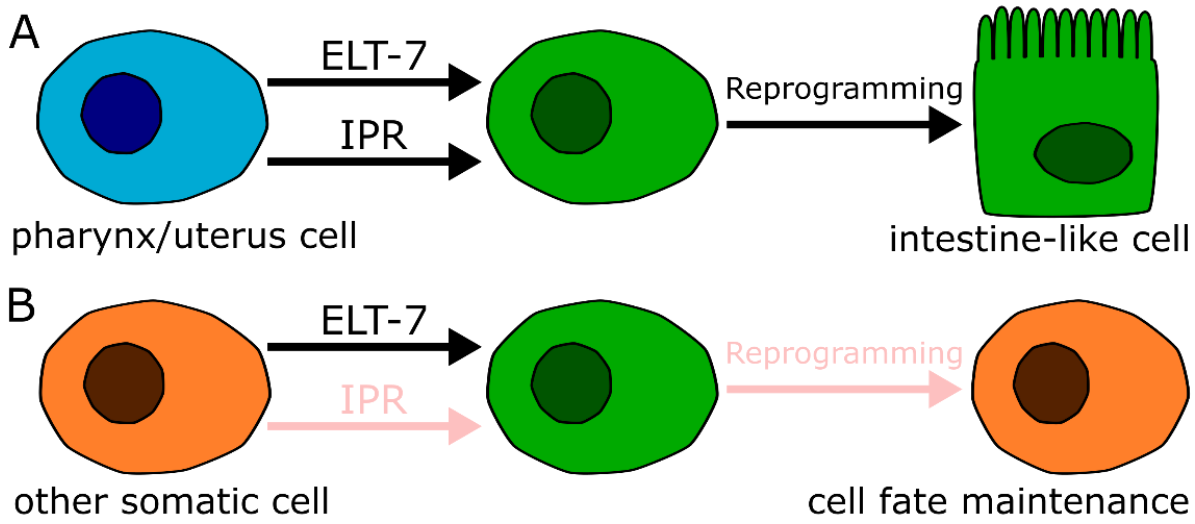

A
ELT-7
IPR
Reprogramming
pharynx/uterus cell
intestine-like cell
B
ELT-7
IPR
Reprogramming
other somatic cell
cell fate maintenance

Fig 7. Model for the role of innate immunity in promoting *in vivo* cellular reprogramming.

(A) Within the context of pharynx and uterus tissues in L4 worms, induced ectopic expression of ELT-7 coincides with the activation of an intracellular pathogen response (IPR) resulting in expression of intestine-specific genes, downregulation of pharyngeal/uterine genes, and results in reprogramming to an intestine cell fate. (B) In other somatic cell types ELT-7 is similarly ectopically expressed, some intestine-specific genes are activated, and tissue-specific genes are transiently downregulated. However, no IPR activation is observed, reprogramming fails, and the transcriptional state of the initial cell fate is maintained, suggesting a role for innate immunity activation in promoting ELT-7-driven reprogramming to an intestine cell fate.

Fig S1.

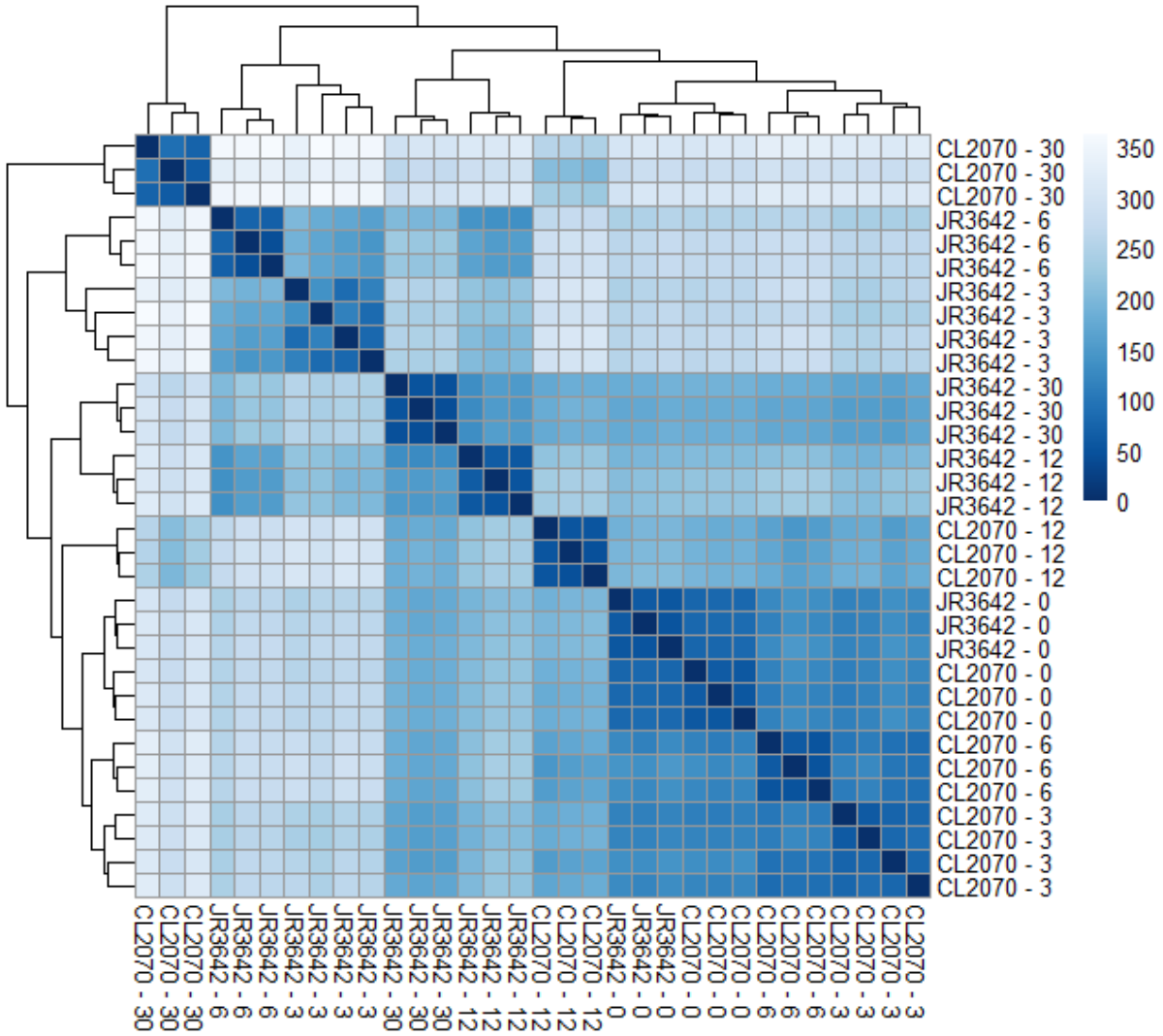

CL2070 - 30
CL2070 - 30
CL2070 - 30
JR3642 - 6
JR3642 - 6
JR3642 - 6
JR3642 - 3
JR3642 - 3
JR3642 - 3
JR3642 - 3
JR3642 - 30
JR3642 - 30
JR3642 - 30
JR3642 - 12
JR3642 - 12
JR3642 - 12
CL2070 - 12
CL2070 - 12
CL2070 - 12
JR3642 - 0
JR3642 - 0
JR3642 - 0
CL2070 - 0
CL2070 - 0
CL2070 - 0
CL2070 - 6
CL2070 - 6
CL2070 - 6
CL2070 - 3
CL2070 - 3
CL2070 - 3
CL2070 - 3
350
300
250
200
150
100
50
0

Fig S1. mRNA-seq time course sample distances.

Regularized logarithm transformed count data from DESeq2 for all 32 mRNA-seq samples was used to calculate similarity between samples. For all conditions, biological replicate samples share the greatest similarity and all 0 hrs PHS samples are highly similar. The most dissimilar samples are 30 hrs PHS controls, which may be attributable to embryonic gene transcription. Smaller values (dark blue) indicate greater similarity. Labels indicate strain – hours post heat shock. JR3642, *hsp::elt-7*; CL2070, *hsp::gfp.*

Fig S2.

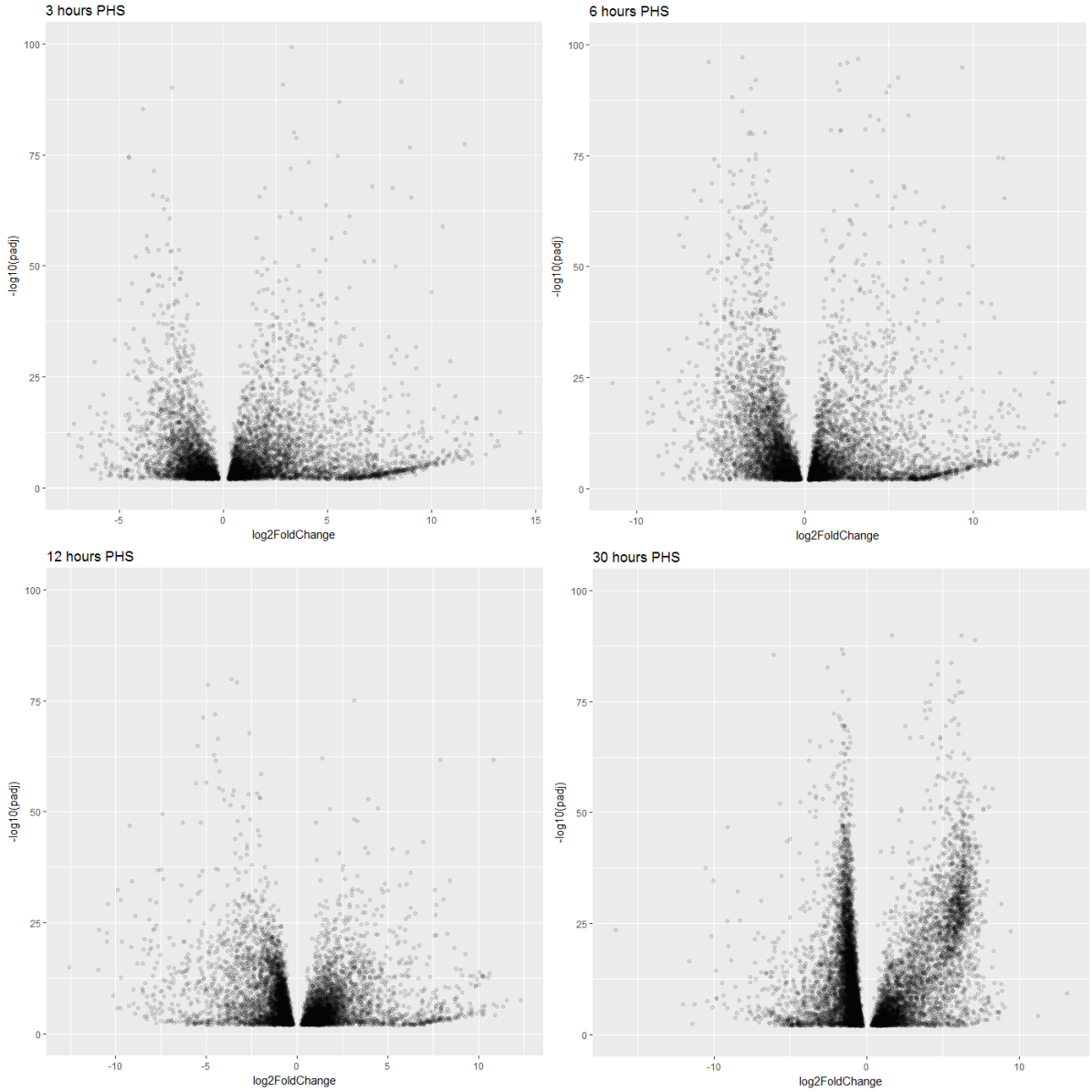

3 hours PHS
6 hours PHS
12 hours PHS
30 hours PHS
-log10(padj)
log2FoldChange

Fig S2. Volcano plots of differentially expressed genes at four time points post *elt-7* induction.

Differential gene expression was calculated using DESeq2 with an FDR of 0.01 for *hsp::elt-7* versus *hsp::gfp* samples. The $\log_2$ fold change and $-\log_{10}$ adjusted p-values are shown for DEGs at 3, 6, 12, and 30 hrs PHS. Each dot represents a gene. The majority of DEGs have relatively small fold changes and larger p-values, although fold change values range from over 1,000-fold downregulated to over 1,000-fold upregulated, and some DEGs have adjusted p-values less than 10^{-90}. There is a notable cluster of DEGs that are strongly upregulated at 3, 6, and 12 hrs PHS with larger p-values and may represent genes that are strongly but variably upregulated following ectopic *elt-7* expression. At 30 hrs PHS there are more significantly downregulated DEGs but with less dramatic fold changes.

Fig S3.

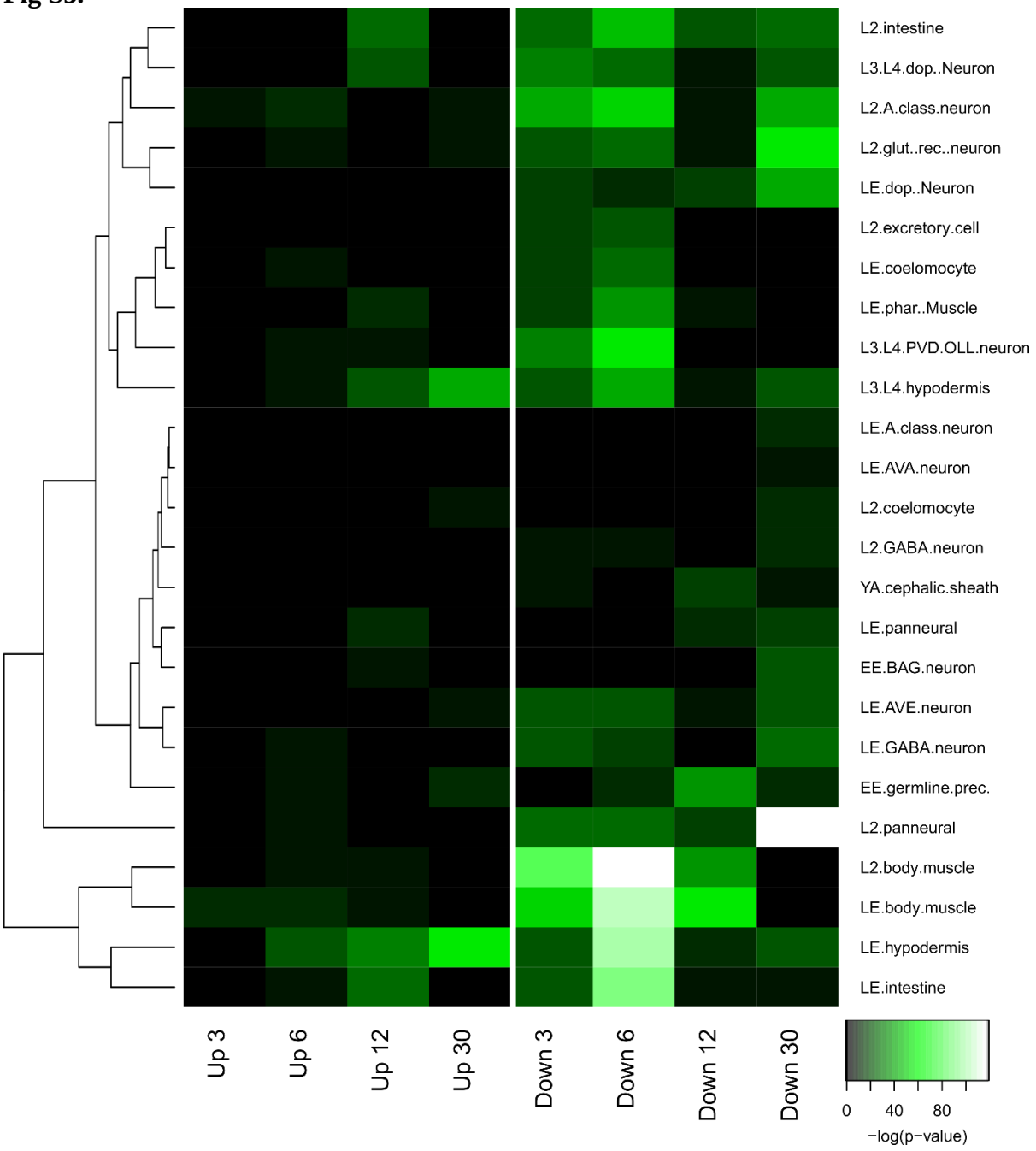

L2.intestine
L3.L4.dop..Neuron
L2.A.class.neuron
L2.glut..rec..neuron
LE.dop..Neuron
L2.excretory.cell
LE.coelomocyte
LE.phar..Muscle
L3.L4.PVD.OLL.neuron
L3.L4.hypodermis
LE.A.class.neuron
LE.AVA.neuron
L2.coelomocyte
L2.GABA.neuron
YA.cephalic.sheath
LE.panneural
EE.BAG.neuron
LE.AVE.neuron
LE.GABA.neuron
EE.germline.prec.
L2.panneural
L2.body.muscle
LE.body.muscle
LE.hypodermis
LE.intestine
Up 3
Up 6
Up 12
Up 30
Down 3
Down 6
Down 12
Down 30
0
40
80
−log(p−value)

Fig S3. Correlation of DEGs and tissue type-specific gene expression.

Tissue-enriched gene sets were cross compared to up- and downregulated DEGs sets and Fisher's exact test was applied to determine significant enrichment. Heatmap displays -log p-values for all tissue-DEG comparisons, darker shades indicate weaker correlations and lighter green shades indicate stronger correlations. Rows are clustered by similarity across all DEG sets. Tissues in the middle cluster have relatively stable tissue specific gene expression following ectopic *elt-7* expression. Tissues in the top and bottom clusters have a similar trend with transient downregulation of tissue specific gene expression. Interestingly, tissue specific genes are more strongly associated with downregulated DEGs at 3 and 6 hrs PHS despite there being more upregulated DEGs at these time points. This suggests that ELT-7 is not disrupting gene expression stochastically but rather specifically inhibiting genes specific to differentiated tissues and upregulating more general response genes.

Fig S4.

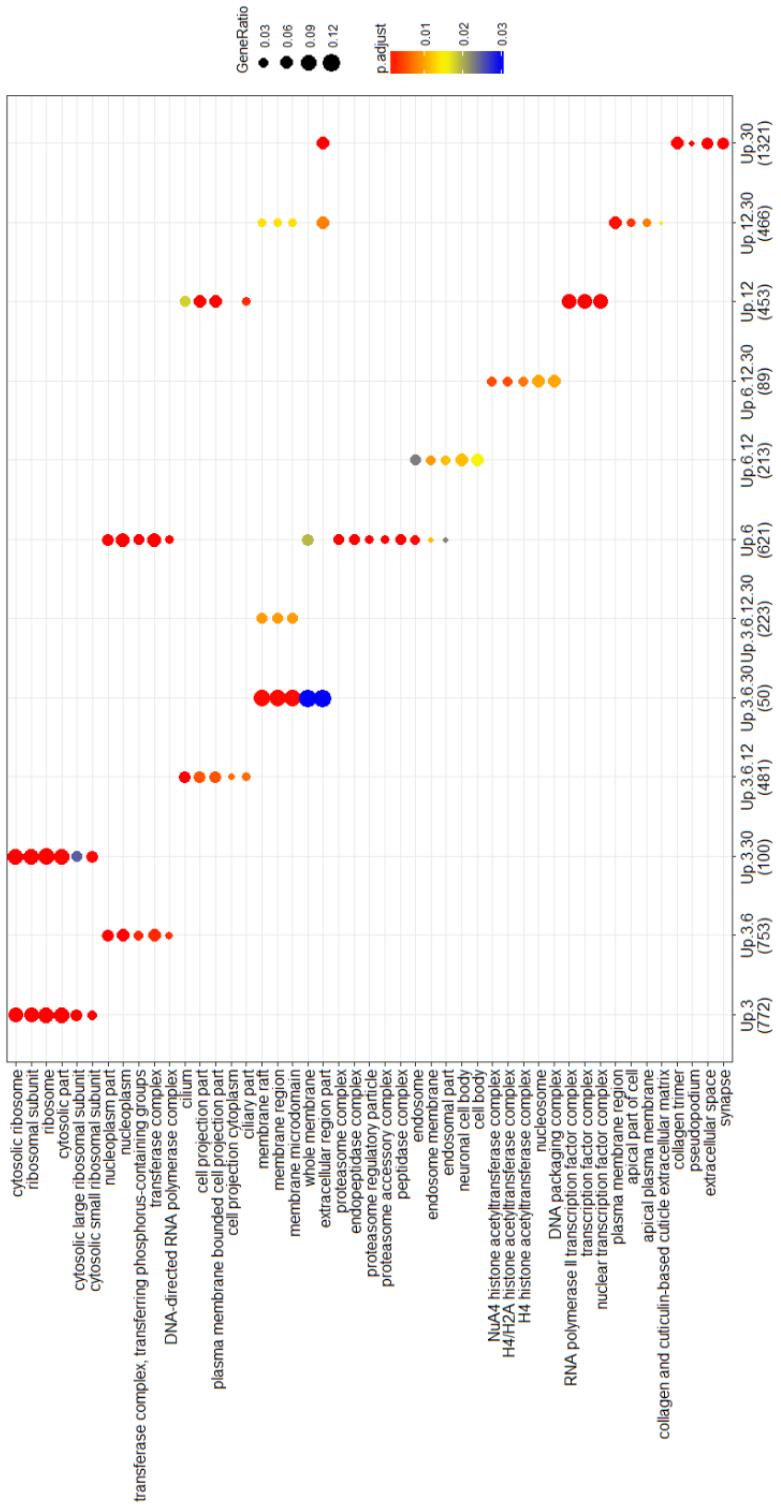

GeneRatio
0.03
0.06
0.09
0.12
p.adjust
0.01
0.02
0.03
Up.3 (772)
Up.3.6 (753)
Up.3.30 (100)
Up.3.6.12 (481)
Up.3.6.30 (50)
Up.3.6.12.30 (223)
Up.6 (621)
Up.6.12 (213)
Up.6.12.30 (89)
Up.12 (453)
Up.12.30 (466)
Up.30 (1321)
cytosolic ribosome
ribosomal subunit
ribosome
cytosolic part
cytosolic large ribosomal subunit
cytosolic small ribosomal subunit
nucleoplasm part
nucleoplasm
transferase complex, transferring phosphorus-containing groups
transferase complex
DNA-directed RNA polymerase complex
cilium
cell projection part
plasma membrane bounded cell projection part
cell projection cytoplasm
ciliary part
membrane raft
membrane region
membrane microdomain
whole membrane
extracellular region part
proteasome complex
endopeptidase complex
proteasome regulatory particle
proteasome accessory complex
peptidase complex
endosome
endosome membrane
endosomal part
neuronal cell body
cell body
NuA4 histone acetyltransferase complex
H4/H2A histone acetyltransferase complex
H4 histone acetyltransferase complex
nucleosome
DNA packaging complex
RNA polymerase II transcription factor complex
transcription factor complex
nuclear transcription factor complex
plasma membrane region
apical part of cell
apical plasma membrane
collagen and cuticulin-based cuticle extracellular matrix
collagen trimer
pseudopodium
extracellular space
synapse

Fig S4. Dotplot of the top Cellular Component GO terms enriched for each upregulated DEG cluster.

Dot column labels indicate the specific time points at which the cluster of genes are upregulated and the number of genes with sufficient annotation that were used for GO enrichment analysis in each cluster. Dot size represents the proportion of genes in the cluster annotated with each term. Dot color represents the adjusted p-value with warmer shades indicating greater significance of enrichment. Interestingly, the rough sequence of events regarding upregulated genes starts with translational machinery, then structural and membrane components, followed by transcriptional and chromatin related genes. Clusters not shown in dotplots had no significant GO enrichments.

Fig S5.

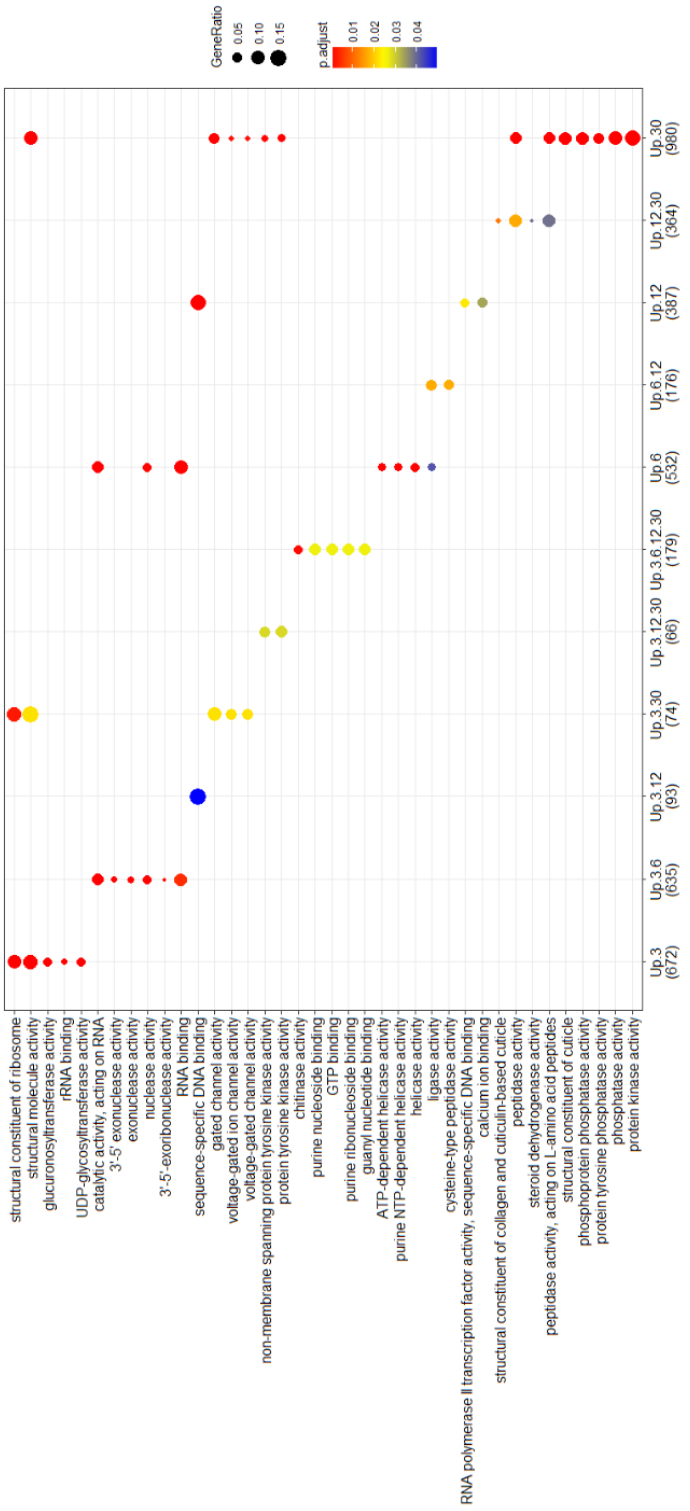

GeneRatio
0.05
0.10
0.15
p.adjust
0.01
0.02
0.03
0.04
Up.3 (672)
Up.3.6 (635)
Up.3.12 (93)
Up.3.30 (74)
Up.3.12.30 (66)
Up.3.6.12.30 (179)
Up.6 (532)
Up.6.12 (176)
Up.12 (387)
Up.12.30 (364)
Up.30 (980)
structural constituent of ribosome
structural molecule activity
glucuronosyltransferase activity
rRNA binding
UDP-glycosyltransferase activity
catalytic activity, acting on RNA
3'-5' exonuclease activity
exonuclease activity
nuclease activity
3'-5'-exoribonuclease activity
RNA binding
sequence-specific DNA binding
gated channel activity
voltage-gated ion channel activity
voltage-gated channel activity
non-membrane spanning protein tyrosine kinase activity
protein tyrosine kinase activity
chitinase activity
purine nucleoside binding
GTP binding
purine ribonucleoside binding
guanyl nucleotide binding
ATP-dependent helicase activity
purine NTP-dependent helicase activity
helicase activity
ligase activity
cysteine-type peptidase activity
RNA polymerase II transcription factor activity, sequence-specific DNA binding
calcium ion binding
structural constituent of collagen and cuticulin-based cuticle
peptidase activity
steroid dehydrogenase activity
peptidase activity, acting on L-amino acid peptides
structural constituent of cuticle
phosphoprotein phosphatase activity
protein tyrosine phosphatase activity
phosphatase activity
protein kinase activity

Fig S5. Dotplot of the top Molecular Function GO terms enriched for each upregulated DEG cluster.

Dot column labels indicate the specific time points at which the cluster of genes are upregulated and the number of genes with sufficient annotation that were used for GO enrichment analysis in each cluster. Dot size represents the proportion of genes in the cluster annotated with each term. Dot color represents the adjusted p-value with warmer shades indicating greater significance of enrichment. Overall there is poor enrichment of molecular function GO terms, indicating that ectopic ELT-7 activity does not specifically upregulate the expression of genes within the same functional class. Clusters not shown in dotplots had no significant GO enrichments.

Fig S6.

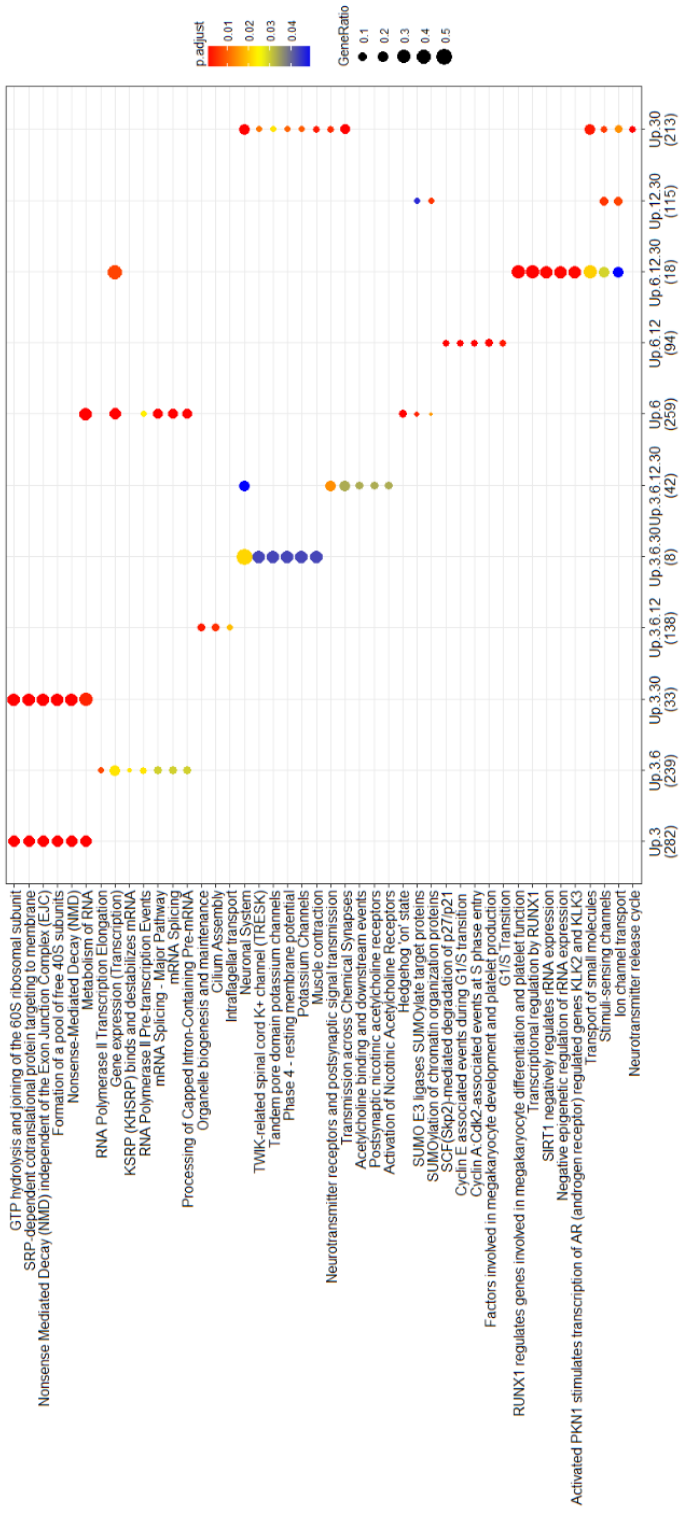

p.adjust
0.01
0.02
0.03
0.04
GeneRatio
0.1
0.2
0.3
0.4
0.5
GTP hydrolysis and joining of the 60S ribosomal subunit
SRP-dependent cotranslational protein targeting to membrane
Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC)
Formation of a pool of free 40S subunits
Nonsense-Mediated Decay (NMD)
Metabolism of RNA
RNA Polymerase II Transcription Elongation
Gene expression (Transcription)
KSRP (KHSRP) binds and destabilizes mRNA
RNA Polymerase II Pre-transcription Events
mRNA Splicing - Major Pathway
mRNA Splicing
Processing of Capped Intron-Containing Pre-mRNA
Organelle biogenesis and maintenance
Cilium Assembly
Intraflagellar transport
Neuronal System
TWIK-related spinal cord K+ channel (TRESK)
Tandem pore domain potassium channels
Phase 4 - resting membrane potential
Potassium Channels
Muscle contraction
Neurotransmitter receptors and postsynaptic signal transmission
Transmission across Chemical Synapses
Acetylcholine binding and downstream events
Postsynaptic nicotinic acetylcholine receptors
Activation of Nicotinic Acetylcholine Receptors
Hedgehog 'on' state
SUMO E3 ligases SUMOylate target proteins
SUMOylation of chromatin organization proteins
SCF(Skp2)-mediated degradation of p27/p21
Cyclin E associated events during G1/S transition
Cyclin A:Cdk2-associated events at S phase entry
Factors involved in megakaryocyte development and platelet production
G1/S Transition
RUNX1 regulates genes involved in megakaryocyte differentiation and platelet function
Transcriptional regulation by RUNX1
SIRT1 negatively regulates rRNA expression
Negative epigenetic regulation of rRNA expression
Activated PKN1 stimulates transcription of AR (androgen receptor) regulated genes KLK2 and KLK3
Transport of small molecules
Stimuli-sensing channels
Ion channel transport
Neurotransmitter release cycle
Up.3 (282)
Up.3.6 (239)
Up.3.30 (33)
Up.3.6.12 (138)
Up.3.6.30 (8)
Up.3.6.12.30 (42)
Up.6 (259)
Up.6.12 (94)
Up.6.12.30 (18)
Up.12.30 (115)
Up.30 (213)

Fig S6. Dotplot of the top Reactome Pathway terms enriched for each upregulated DEG cluster.

Dot column labels indicate the specific time points at which the cluster of genes are upregulated and the number of genes with sufficient annotation that were used for Reactome pathway term enrichment analysis in each cluster. Dot size represents the proportion of genes in the cluster annotated with each term. Dot color represents the adjusted p-value with warmer shades indicating greater significance of enrichment. Most notably are upregulated pathway terms related to RNA processing and translation at early time points, which may be attributable to the burst in additional transcription factor production (ELT-7) and its transcription promoting activity. Clusters not shown in dotplots had no significant GO enrichments.

Fig S7.

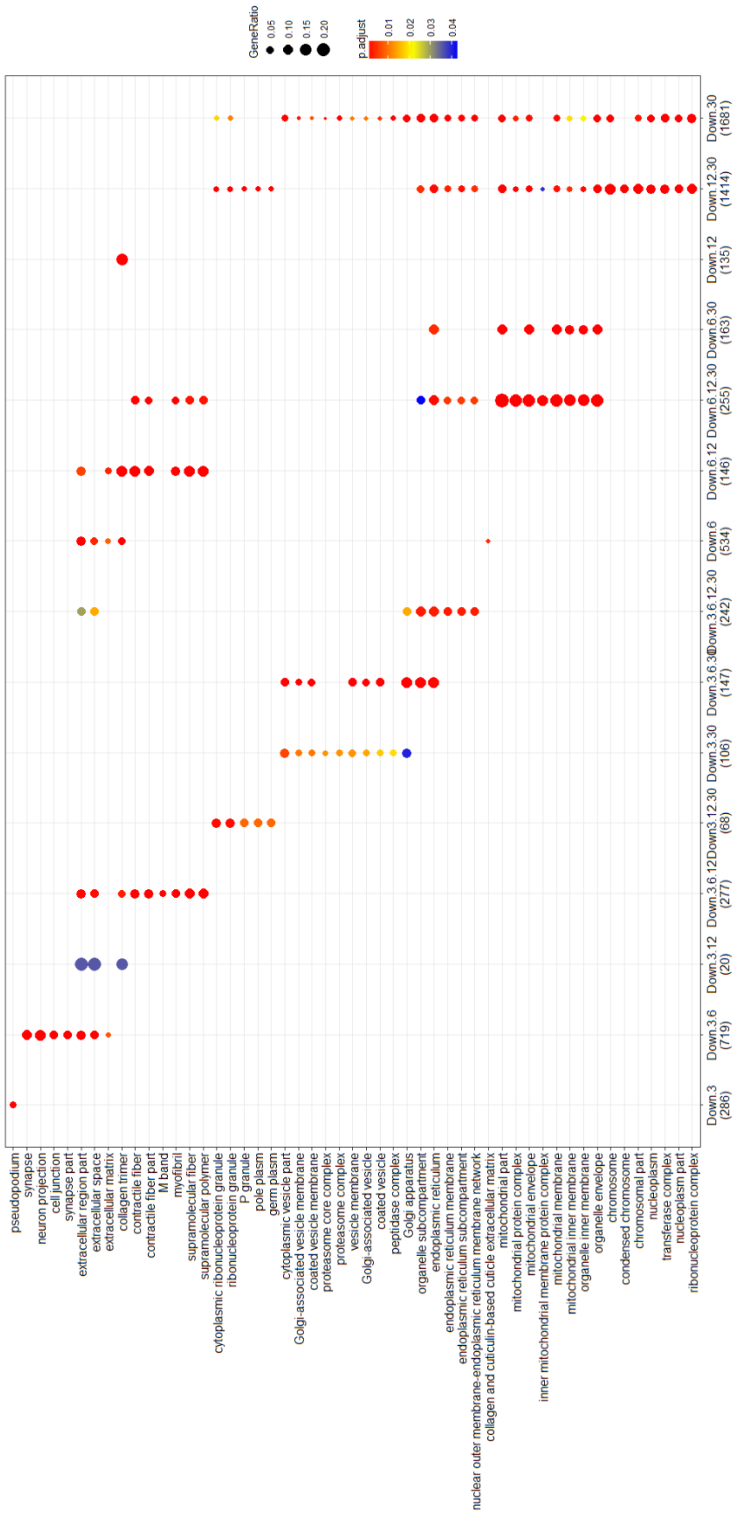

GeneRatio
0.05
0.10
0.15
0.20
p.adjust
0.01
0.02
0.03
0.04
Down.3 (286)
Down.3.6 (719)
Down.3.12 (20)
Down.3.6.12 (277)
Down.3.12.30 (68)
Down.3.30 (106)
Down.3.6.30 (147)
Down.3.6.12.30 (242)
Down.6 (534)
Down.6.12 (146)
Down.6.12.30 (255)
Down.6.30 (163)
Down.12 (135)
Down.12.30 (1414)
Down.30 (1681)
pseudopodium
synapse
neuron projection
cell junction
synapse part
extracellular region part
extracellular space
extracellular matrix
collagen trimer
contractile fiber
contractile fiber part
M band
myofibril
supramolecular fiber
supramolecular polymer
cytoplasmic ribonucleoprotein granule
ribonucleoprotein granule
P granule
pole plasm
germ plasm
cytoplasmic vesicle part
Golgi-associated vesicle membrane
coated vesicle membrane
proteasome core complex
proteasome complex
vesicle membrane
Golgi-associated vesicle
coated vesicle
peptidase complex
Golgi apparatus
organelle subcompartment
endoplasmic reticulum
endoplasmic reticulum membrane
endoplasmic reticulum subcompartment
nuclear outer membrane-endoplasmic reticulum membrane network
collagen and cuticulin-based cuticle extracellular matrix
mitochondrial part
mitochondrial protein complex
mitochondrial envelope
inner mitochondrial membrane protein complex
mitochondrial membrane
mitochondrial inner membrane
organelle inner membrane
organelle envelope
chromosome
condensed chromosome
chromosomal part
nucleoplasm
transferase complex
nucleoplasm part
ribonucleoprotein complex

Fig S7. Dotplot of the top Cellular Component GO terms enriched for each downregulated DEG cluster.

Dot column labels indicate the specific time points at which the cluster of genes are downregulated and the number of genes with sufficient annotation that were used for GO enrichment analysis in each cluster. Dot size represents the proportion of genes in the cluster annotated with each term. Dot color represents the adjusted p-value with warmer shades indicating greater significance of enrichment. Of note are the multiple cellular component GO terms related to muscle function (ie: “contractile fiber” and “myofibril”) that are enriched for genes downregulated 3, 6, and 12 hrs PHS, which further supports the observed transient downregulation of tissue-specific genes in non-reprogrammed cell types. Clusters not shown in dotplots had no significant GO enrichments.

Fig S8.

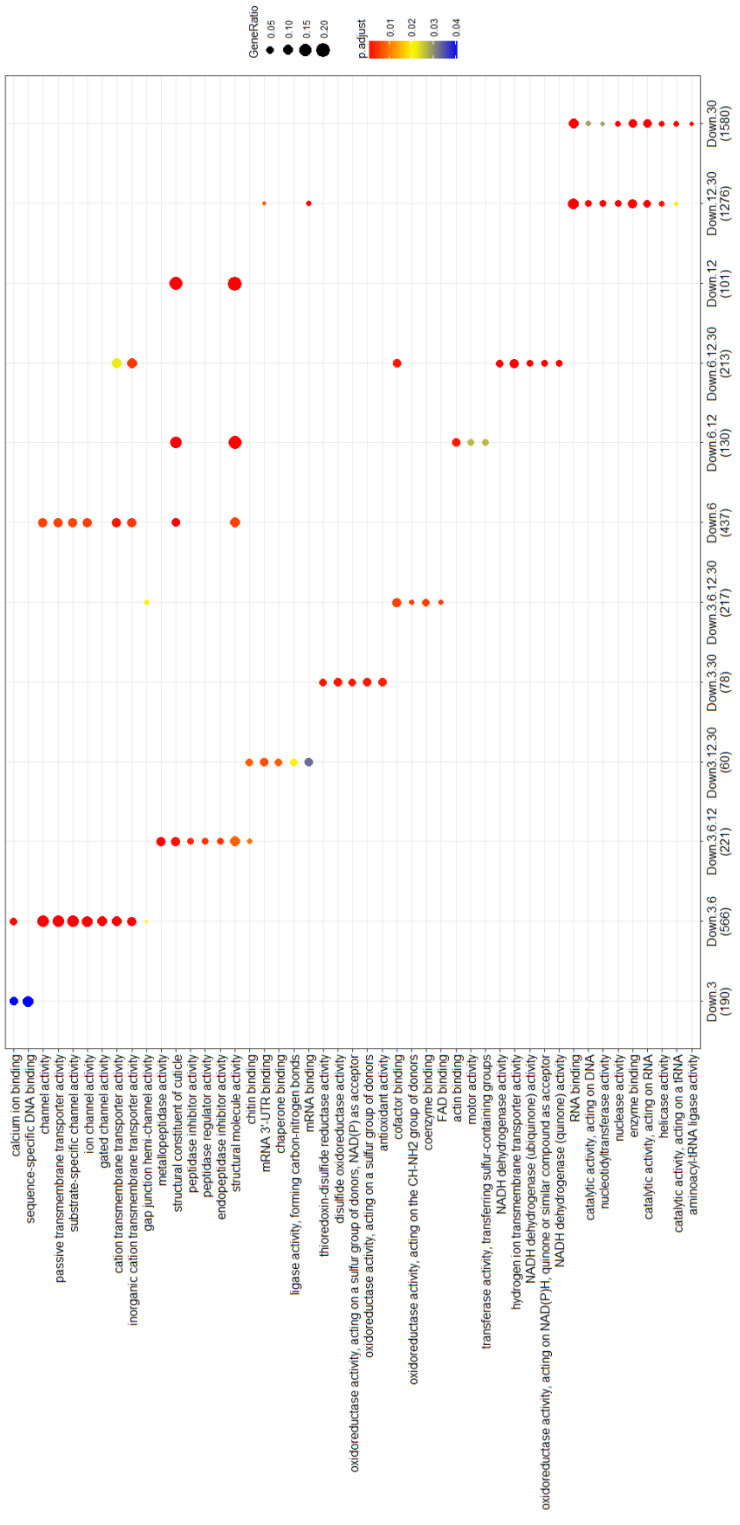

GeneRatio
0.05
0.10
0.15
0.20
p.adjust
0.01
0.02
0.03
0.04
Down.3 (190)
Down.3.6 (566)
Down.3.6.12 (221)
Down.3.12.30 (60)
Down.3.30 (78)
Down.3.6.12.30 (217)
Down.6 (437)
Down.6.12 (130)
Down.6.12.30 (213)
Down.12 (101)
Down.12.30 (1276)
Down.30 (1580)
calcium ion binding
sequence-specific DNA binding
channel activity
passive transmembrane transporter activity
substrate-specific channel activity
ion channel activity
gated channel activity
cation transmembrane transporter activity
inorganic cation transmembrane transporter activity
gap junction hemi-channel activity
metallopeptidase activity
structural constituent of cuticle
peptidase inhibitor activity
peptidase regulator activity
endopeptidase inhibitor activity
structural molecule activity
chitin binding
mRNA 3'-UTR binding
chaperone binding
ligase activity, forming carbon-nitrogen bonds
mRNA binding
thioredoxin-disulfide reductase activity
disulfide oxidoreductase activity
oxidoreductase activity, acting on a sulfur group of donors; NAD(P) as acceptor
oxidoreductase activity, acting on a sulfur group of donors
antioxidant activity
cofactor binding
oxidoreductase activity, acting on the CH-NH2 group of donors
coenzyme binding
FAD binding
actin binding
motor activity
transferase activity, transferring sulfur-containing groups
NADH dehydrogenase activity
hydrogen ion transmembrane transporter activity
NADH dehydrogenase (ubiquinone) activity
oxidoreductase activity, acting on NAD(P)H, quinone or similar compound as acceptor
NADH dehydrogenase (quinone) activity
RNA binding
catalytic activity, acting on DNA
nucleotidyltransferase activity
nuclease activity
enzyme binding
catalytic activity, acting on RNA
helicase activity
catalytic activity, acting on a tRNA
aminoacyl-tRNA ligase activity

Fig S8. Dotplot of the top Molecular Function GO terms enriched for each downregulated DEG cluster.

Dot column labels indicate the specific time points at which the cluster of genes are downregulated and the number of genes with sufficient annotation that were used for GO enrichment analysis in each cluster. Dot size represents the proportion of genes in the cluster annotated with each term. Dot color represents the adjusted p-value with warmer shades indicating greater significance of enrichment. Numerous molecular function terms related to oxidative reduction and electron transport are downregulated at 6, 12, and 30 hrs PHS, which may be attributable to a collapse of mitochondrial functioning as a downstream consequence of global ectopic transcription factor expression and forced cellular reprogramming. Clusters not shown in dotplots had no significant GO enrichments.

Fig S9.

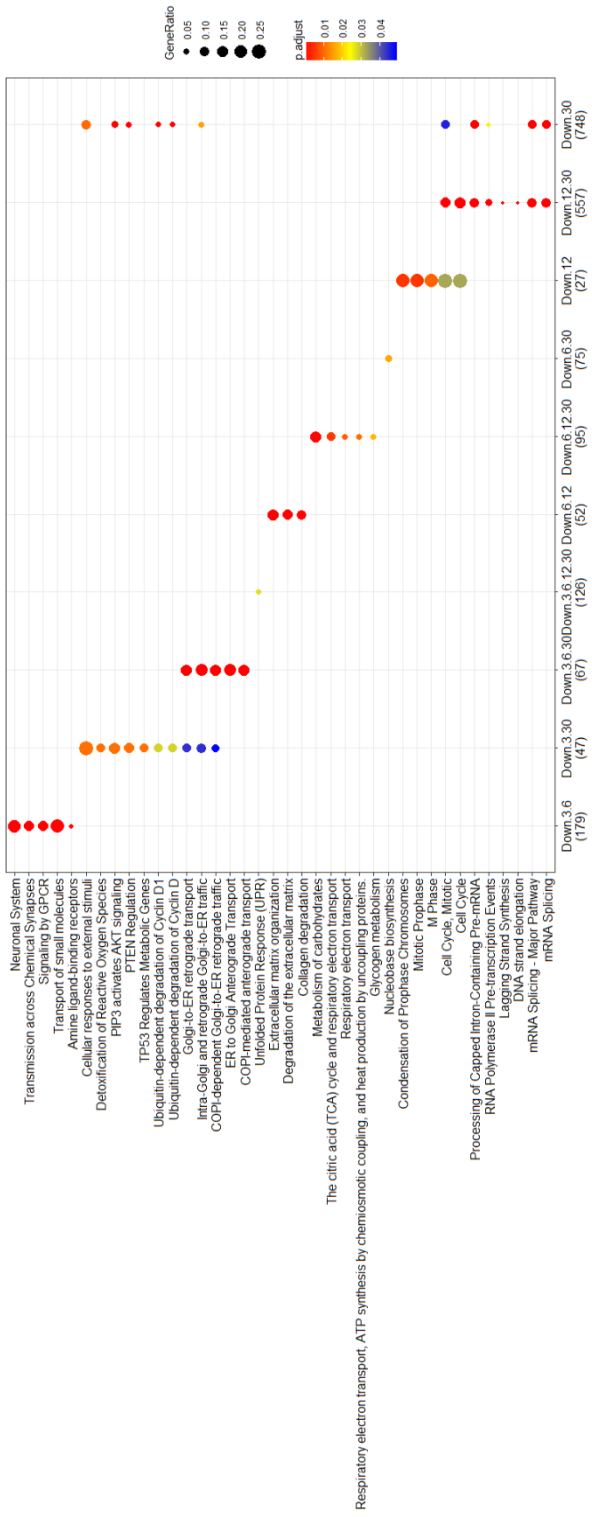

GeneRatio
0.05
0.10
0.15
0.20
0.25
p.adjust
0.01
0.02
0.03
0.04
Down.3.6 (179)
Down.3.30 (47)
Down.3.6.30 (67)
Down.3.6.12.30 (126)
Down.6.12 (52)
Down.6.12.30 (95)
Down.6.30 (75)
Down.12 (27)
Down.12.30 (557)
Down.30 (748)
Neuronal System
Transmission across Chemical Synapses
Signaling by GPCR
Transport of small molecules
Amine ligand-binding receptors
Cellular responses to external stimuli
Detoxification of Reactive Oxygen Species
PIP3 activates AKT signaling
PTEN Regulation
TP53 Regulates Metabolic Genes
Ubiquitin-dependent degradation of Cyclin D1
Ubiquitin-dependent degradation of Cyclin D
Golgi-to-ER retrograde transport
Intra-Golgi and retrograde Golgi-to-ER traffic
COPI-dependent Golgi-to-ER retrograde traffic
ER to Golgi Anterograde Transport
COPI-mediated anterograde transport
Unfolded Protein Response (UPR)
Extracellular matrix organization
Degradation of the extracellular matrix
Collagen degradation
Metabolism of carbohydrates
The citric acid (TCA) cycle and respiratory electron transport
Respiratory electron transport
Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins
Glycogen metabolism
Nucleobase biosynthesis
Condensation of Prophase Chromosomes
Mitotic Prophase
M Phase
Cell Cycle, Mitotic
Cell Cycle
Processing of Capped Intron-Containing Pre-mRNA
RNA Polymerase II Pre-transcription Events
Lagging Strand Synthesis
DNA strand elongation
mRNA Splicing - Major Pathway
mRNA Splicing

Fig S9. Dotplot of the top Reactome Pathway terms enriched for each downregulated DEG cluster.

Dot column labels indicate the specific time points at which the cluster of genes are downregulated and the number of genes with sufficient annotation that were used for Reactome pathway term enrichment analysis in each cluster. Dot size represents the proportion of genes in the cluster annotated with each term. Dot color represents the adjusted p-value with warmer shades indicating greater significance of enrichment. The most significantly enriched pathway terms are associated with nervous system functioning at 3 and 6 hrs PHS. It is unclear whether neurons are particularly sensitive to initial ectopic ELT-7 activity or if changes in neuron function are due to another mechanism such as initiation of an aversive learning behavioral response. Clusters not shown in dotplots had no significant GO enrichments.

Fig S10.

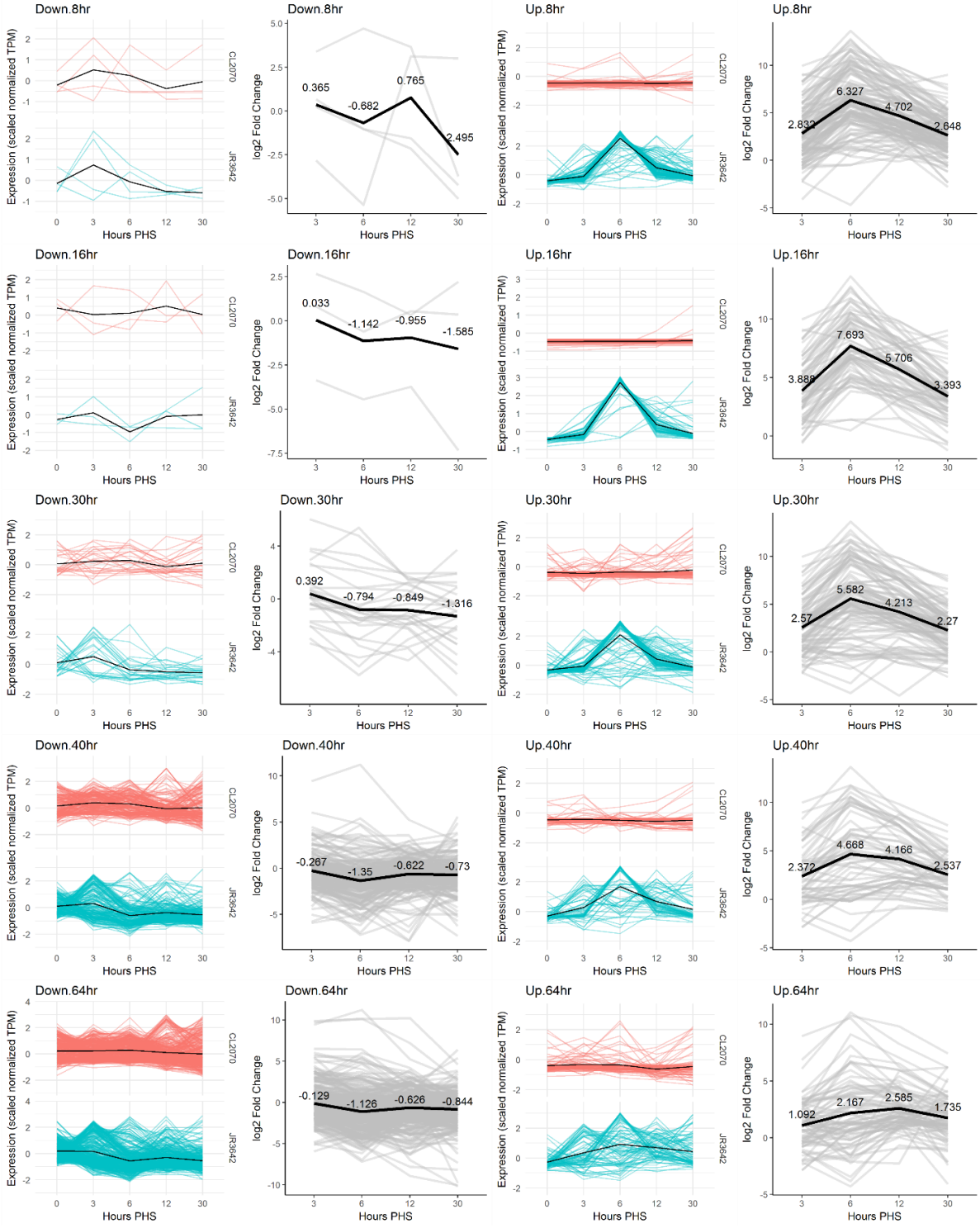

Down.8hr
Up.8hr
Down.16hr
Up.16hr
Down.30hr
Up.30hr
Down.40hr
Up.40hr
Down.64hr
Up.64hr
Expression (scaled normalized TPM)
log2 Fold Change
Hours PHS
CL2070
JR3642
0.365
-0.682
0.765
2.495
6.327
4.702
2.832
2.648
0.033
-1.142
-0.955
-1.585
7.693
5.706
3.888
3.393
0.392
-0.794
-0.849
-1.316
5.582
4.213
2.57
2.27
-0.267
-1.35
-0.622
-0.73
4.668
4.166
2.372
2.537
-0.129
-1.126
-0.626
-0.844
2.167
2.585
1.092
1.735

Fig S10. Ectopic ELT-7 induced gene expression patterns for select genes that are differentially expressed following *N. parisii* infection.

A previous study profiling the temporal transcriptional response of *C. elegans* to intracellular infection by *N. parisii* detected a unique innate immune response (Bakowski et al., 2014). Gene expression patterns and $\log_2$ fold changes following *hsp::elt-7* induction were plotted for up- and downregulated DEGs at each post microsporidia infection time point profiled by Bakowski et al. (8, 16, 30, 40, and 64 hours post infection). Post-*elt-7* gene expression is plotted as scaled, normalized counts for *hsp::gfp* controls (CL2070, red) and *hsp::elt-7* experimental (JR3642, blue) data separately; each colored line corresponds to a gene, black line shows average gene expression level. The differences in *hsp::elt-7* versus *hsp::gfp* gene expression were also plotted as $\log_2$ fold changes for the same groups of genes; each gray line corresponds to one gene, black line shows average $\log_2$ fold change. Genes downregulated following intracellular pathogen infection tend to be downregulated following ectopic *elt-7* expression as well. Genes upregulated following intracellular pathogen infection are strongly and transiently upregulated on average following ectopic *elt-7* expression, with some genes increasing over 1,000-fold.

Table S1. DEG-TFBS correlation raw p-values for 217 TFs.

TF	Up 3	Up 6	Up 12	Up 30	Down 3	Down 6	Down 12	Down 30
AHA-1	8.80E-01	3.26E-02	8.73E-04	1.83E-32	6.75E-02	6.80E-03	4.25E-16	4.35E-59
AHR-1	2.36E-01	7.88E-03	4.94E-09	1.07E-19	9.54E-05	7.55E-03	1.87E-29	4.06E-53
ALR-1	2.33E-01	2.16E-02	8.75E-01	4.34E-06	7.35E-01	5.42E-01	9.81E-03	2.16E-08
ATF-7	3.18E-02	6.03E-05	4.71E-01	3.96E-25	8.11E-02	4.12E-01	2.84E-06	1.79E-40
ATHP-1	3.11E-01	4.41E-07	5.39E-14	1.56E-52	8.60E-08	5.23E-11	5.59E-49	6.68E-120
B0035.1	6.51E-01	1.08E-03	2.81E-02	4.24E-17	3.75E-03	9.65E-03	7.13E-15	4.37E-36
B0261.1	7.13E-01	9.14E-03	4.72E-02	7.76E-15	9.47E-02	5.46E-02	6.71E-10	1.80E-23
B0310.2	2.65E-01	7.83E-02	3.81E-01	3.02E-14	9.62E-02	7.12E-03	2.18E-08	4.17E-23
BLMP-1	7.97E-01	5.43E-07	1.06E-01	6.36E-65	3.43E-01	9.65E-01	9.23E-57	7.18E-252
C01B12.2	7.01E-02	5.68E-19	4.38E-04	1.18E-119	1.78E-16	1.95E-14	3.71E-90	0.00E+00
C04F5.9	1.00E+00	2.99E-08	3.71E-20	8.51E-90	3.88E-08	3.74E-12	1.72E-73	1.91E-225
C05D10.1	5.19E-01	8.97E-01	5.01E-01	8.18E-04	4.67E-01	4.34E-01	7.09E-02	2.74E-06
C08G9.2	1.03E-01	2.02E-04	5.66E-07	1.10E-41	5.35E-04	9.53E-05	2.19E-48	5.10E-120
C16A3.4	6.47E-01	4.07E-18	1.23E-14	4.52E-159	3.33E-12	1.81E-17	1.97E-106	0.00E+00
C27D6.4	7.98E-01	2.04E-12	1.44E-01	1.19E-116	6.50E-02	3.46E-01	1.75E-51	2.21E-253
C34B4.2	8.77E-01	4.33E-01	3.27E-04	1.81E-14	2.17E-01	4.32E-01	1.14E-17	5.54E-29
C34F6.9	5.81E-02	8.98E-21	1.55E-10	9.52E-181	2.49E-03	7.74E-06	7.79E-133	0.00E+00
CEBP-1	6.85E-03	1.42E-39	3.72E-01	1.81E-105	8.71E-07	1.52E-12	5.06E-42	1.58E-232
CEH-14	2.30E-02	3.26E-05	4.17E-06	5.25E-47	5.01E-03	3.15E-04	1.47E-43	8.01E-125
CEH-16	9.68E-01	3.75E-01	1.12E-01	1.47E-30	5.68E-03	3.49E-03	3.25E-16	4.51E-70
CEH-18	2.14E-03	1.57E-03	8.02E-08	9.11E-63	3.81E-02	8.97E-02	2.53E-59	2.20E-139
CEH-2	2.31E-01	1.88E-06	1.62E-05	1.00E-53	1.90E-06	1.70E-06	1.18E-48	2.01E-134
CEH-22	1.21E-01	2.98E-01	2.34E-01	5.63E-07	6.39E-01	1.00E+00	2.45E-03	3.57E-10
CEH-24	3.00E-01	9.24E-01	2.83E-03	1.13E-07	1.63E-01	7.02E-01	5.12E-10	8.78E-16
CEH-26	7.27E-01	2.95E-02	8.55E-01	1.12E-16	1.00E+00	1.30E-02	8.43E-08	9.09E-26
CEH-30	3.64E-01	2.38E-04	6.51E-03	1.02E-20	1.24E-02	5.10E-06	1.07E-10	9.82E-39
CEH-31	4.27E-02	1.76E-08	2.43E-07	3.62E-114	3.68E-03	2.12E-06	1.48E-72	2.10E-251
CEH-32	2.38E-01	2.32E-01	7.41E-01	7.96E-07	5.92E-01	3.37E-01	3.62E-06	1.94E-15
CEH-34	6.00E-02	1.20E-05	1.72E-04	9.47E-58	8.36E-02	1.30E-02	2.69E-38	9.89E-120
CEH-36	1.24E-01	1.50E-01	1.52E-01	6.22E-09	3.66E-01	9.08E-01	4.29E-11	1.73E-23
CEH-38	3.69E-01	6.98E-07	9.00E-01	7.27E-60	9.87E-04	1.09E-08	4.59E-25	7.71E-159
CEH-39	1.00E+00	1.74E-16	4.08E-14	6.58E-130	2.54E-16	1.28E-26	2.80E-66	0.00E+00
CEH-48	7.33E-03	4.10E-05	2.07E-02	3.31E-59	7.57E-03	5.90E-09	1.03E-27	1.75E-135
CEH-79	1.97E-01	1.11E-02	2.21E-02	3.54E-10	3.23E-03	7.11E-02	1.48E-03	1.48E-11
CEH-9	6.12E-01	6.08E-01	3.96E-01	4.76E-03	6.46E-01	7.58E-01	2.92E-03	7.15E-04
CEH-90	4.17E-01	1.01E-06	6.31E-13	2.87E-49	1.51E-08	2.05E-08	9.62E-51	4.68E-127
CES-1	3.23E-02	2.27E-10	8.17E-06	5.33E-110	8.86E-01	8.19E-04	3.15E-65	1.57E-280
CEY-2	6.94E-01	2.03E-01	8.69E-01	2.25E-05	4.75E-01	3.80E-01	8.74E-01	7.17E-03
CHD-7	7.65E-01	7.38E-04	1.94E-01	8.01E-35	7.34E-01	3.73E-02	3.82E-10	2.98E-50
CHE-1	1.44E-01	1.63E-02	3.42E-04	3.11E-18	3.80E-02	1.22E-01	1.88E-16	3.08E-43
COG-1	3.09E-01	7.08E-07	5.87E-04	2.18E-66	1.33E-03	1.52E-05	7.90E-36	5.30E-134
DAF-12	2.82E-01	5.90E-01	2.60E-01	1.19E-03	4.50E-01	5.70E-02	5.00E-07	2.60E-08
DAF-16	6.97E-01	2.39E-11	5.00E-10	3.77E-105	6.58E-05	6.66E-17	1.96E-29	1.19E-171
DIE-1	5.07E-01	1.47E-01	4.53E-03	6.56E-15	8.02E-01	1.00E+00	8.40E-11	7.37E-22
DMD-4	9.27E-02	8.65E-03	2.18E-02	3.36E-24	6.45E-02	1.60E-03	3.75E-20	3.32E-53
DPFF-1	1.10E-01	2.31E-03	2.00E-01	9.42E-07	3.24E-02	1.05E-01	1.06E-02	3.18E-09
DPL-1	2.14E-01	3.88E-22	6.93E-18	3.26E-226	1.35E-14	1.57E-29	2.71E-200	0.00E+00
DSC-1	1.19E-01	1.30E-05	4.39E-11	1.76E-94	3.40E-01	6.60E-01	1.41E-77	3.99E-225
DUXL-1	4.50E-01	7.37E-07	1.54E-03	4.41E-55	6.22E-05	6.83E-04	1.11E-30	1.56E-108
DVE-1	8.29E-01	6.20E-11	1.61E-04	2.38E-111	4.20E-01	2.38E-04	7.85E-25	5.81E-186
EFL-1	1.91E-01	1.98E-23	1.34E-29	2.09E-264	7.25E-15	1.24E-28	3.24E-247	0.00E+00
EGL-13	7.96E-01	1.35E-07	2.52E-08	1.28E-76	8.45E-05	3.96E-08	1.25E-52	1.57E-179
EGL-27	2.60E-01	2.87E-04	1.05E-01	2.50E-46	1.86E-02	3.44E-03	1.43E-23	3.50E-85
EGL-5	2.85E-01	1.05E-06	1.17E-05	3.31E-66	1.74E-01	1.19E-04	4.95E-44	9.70E-163
ELT-1	1.16E-04	4.53E-02	8.66E-04	3.43E-13	2.89E-01	3.83E-01	4.06E-10	2.70E-41

TF	Up 3	Up 6	Up 12	Up 30	Down 3	Down 6	Down 12	Down 30
ELT-2	7.96E-09	1.38E-23	8.29E-23	1.01E-81	5.66E-01	1.88E-04	6.32E-17	7.68E-146
ELT-3	1.81E-06	1.68E-08	2.50E-01	1.04E-56	3.16E-12	1.72E-06	6.30E-30	2.83E-129
ELT-4	3.66E-01	6.97E-01	8.93E-01	7.13E-01	3.07E-01	4.34E-01	2.69E-02	1.82E-02
EOR-1	3.83E-01	1.68E-04	4.77E-01	6.45E-10	1.06E-02	2.29E-02	3.09E-02	1.68E-17
ETS-4	4.39E-01	3.10E-07	1.31E-06	3.81E-48	2.30E-06	1.03E-08	4.51E-45	1.51E-123
ETS-7	4.70E-02	1.99E-02	3.11E-03	4.11E-17	5.90E-02	5.95E-02	2.62E-20	4.70E-48
F10B5.3	3.74E-01	3.47E-03	2.69E-02	1.86E-24	4.88E-01	3.20E-02	1.65E-17	5.75E-47
F10E7.11	2.93E-01	6.18E-03	9.26E-03	1.50E-10	8.81E-02	1.09E-01	1.23E-10	4.44E-17
F13C5.2	6.75E-01	6.44E-09	1.08E-09	3.19E-74	1.05E-05	4.62E-07	1.60E-53	1.59E-178
F13H6.1	9.79E-04	8.73E-04	6.90E-06	3.96E-91	1.93E-01	4.90E-02	5.43E-64	4.49E-210
F16B12.6	9.56E-01	4.09E-18	1.26E-25	1.62E-155	2.84E-18	3.24E-27	2.07E-126	0.00E+00
F22D6.2	8.02E-01	1.32E-02	1.14E-03	4.82E-29	3.12E-03	3.73E-02	1.88E-17	1.52E-55
F23B12.7	4.96E-01	2.97E-07	9.15E-26	1.05E-72	7.82E-11	1.62E-12	2.91E-76	3.00E-194
F23F12.9	9.56E-01	3.94E-02	1.66E-01	8.23E-02	6.62E-01	9.55E-01	5.99E-02	5.77E-10
F37D6.2	1.00E+00	4.15E-05	5.85E-07	2.00E-56	4.50E-03	3.83E-06	1.20E-38	3.24E-132
F45C12.2	6.80E-01	8.47E-12	1.71E-16	7.97E-108	2.26E-08	1.56E-12	2.17E-75	3.36E-257
F49E8.2	1.61E-02	1.62E-06	9.20E-39	1.06E-83	3.59E-16	3.63E-24	2.55E-110	4.07E-212
F52B5.7	9.53E-01	8.57E-02	3.80E-03	3.78E-13	4.64E-01	4.75E-01	2.83E-09	2.78E-22
F55B11.4	3.97E-01	1.00E+00	5.55E-01	5.90E-01	9.57E-02	1.49E-01	7.70E-02	3.13E-01
F57C9.4	1.77E-02	3.02E-03	5.97E-02	1.13E-07	1.80E-02	1.63E-02	6.14E-06	8.45E-14
FAX-1	1.32E-01	2.31E-04	1.38E-08	3.23E-34	4.51E-06	3.40E-05	3.83E-44	4.48E-96
FKH-10	9.33E-02	7.81E-04	3.82E-04	6.19E-54	2.82E-01	8.87E-02	7.15E-33	9.71E-111
FKH-2	1.00E+00	7.37E-01	2.96E-01	7.48E-05	2.07E-01	7.36E-01	2.62E-07	8.16E-14
FKH-3	8.65E-02	4.10E-02	3.43E-01	1.87E-05	1.16E-02	1.02E-01	5.37E-02	6.20E-03
FKH-4	5.87E-01	1.00E-02	6.52E-12	6.34E-47	1.05E-02	4.29E-04	2.66E-49	1.29E-105
FKH-6	2.10E-01	4.01E-02	3.21E-03	2.32E-20	1.43E-03	1.62E-03	7.21E-23	6.56E-56
FKH-8	9.37E-01	4.72E-09	1.22E-06	3.94E-98	3.85E-02	2.63E-06	2.84E-47	3.17E-179
FOS-1	2.64E-07	5.88E-20	1.17E-27	1.40E-55	8.10E-01	2.84E-03	4.58E-08	4.79E-103
GMEB-2	7.52E-01	5.89E-03	2.70E-05	1.40E-15	7.77E-03	2.23E-01	9.51E-16	7.74E-39
HIF-1	3.94E-01	2.37E-07	4.57E-06	8.32E-92	3.05E-03	6.78E-03	9.64E-49	3.24E-204
HIM-8	2.85E-01	7.58E-01	4.60E-07	1.47E-26	5.25E-01	5.21E-04	7.58E-39	1.35E-58
HLH-1	3.58E-03	1.59E-01	5.85E-08	1.57E-71	3.54E-04	1.80E-07	1.95E-65	8.90E-154
HLH-12	8.81E-01	7.62E-01	4.07E-02	8.63E-02	6.09E-01	2.89E-01	7.42E-04	8.43E-03
HLH-15	4.17E-01	2.74E-03	6.35E-08	6.20E-13	5.36E-02	2.30E-01	3.28E-16	7.13E-40
HLH-30	1.00E+00	1.74E-11	2.82E-01	3.27E-118	1.00E+00	9.46E-01	2.32E-30	1.33E-219
HLH-4	1.70E-01	6.06E-02	1.40E-02	2.80E-07	1.71E-01	5.13E-01	3.87E-08	2.32E-15
HLH-6	1.01E-01	1.12E-01	9.90E-02	7.30E-15	8.82E-01	2.32E-01	2.38E-15	1.11E-28
HLH-8	3.36E-02	3.00E-05	8.04E-05	9.58E-50	3.72E-02	5.15E-02	3.30E-39	1.21E-104
HMBX-1	5.59E-01	4.27E-02	1.54E-03	6.58E-10	1.78E-02	4.17E-02	7.52E-10	1.71E-23
HMG-11	4.56E-02	7.47E-15	9.24E-06	5.47E-130	9.78E-04	2.11E-07	1.09E-65	0.00E+00
HND-1	7.72E-01	2.31E-03	1.16E-05	3.49E-19	1.97E-03	2.17E-01	2.43E-25	4.02E-47
IRX-1	1.00E+00	1.44E-08	1.11E-11	4.69E-91	2.65E-08	2.94E-07	4.42E-60	6.36E-220
JUN-1	2.15E-03	4.28E-11	1.88E-01	5.74E-54	7.93E-05	1.11E-03	3.99E-22	3.60E-115
K09A11.1	5.19E-01	5.28E-03	4.91E-02	2.01E-20	2.35E-02	1.46E-02	1.59E-17	5.08E-47
LET-607	4.49E-01	1.21E-03	2.09E-02	5.26E-54	2.41E-01	3.11E-01	5.59E-25	1.28E-101
LIM-6	1.59E-01	2.19E-04	3.09E-05	5.51E-24	6.43E-03	3.12E-04	2.81E-29	1.37E-67
LIN-11	5.58E-01	7.59E-06	1.21E-07	1.52E-77	5.62E-05	2.01E-05	2.94E-48	1.83E-171
LIN-13	9.80E-01	9.54E-17	7.16E-11	3.21E-152	1.71E-09	1.42E-15	6.01E-99	0.00E+00
LIN-15B	6.28E-01	4.10E-25	9.21E-22	1.26E-186	2.91E-20	7.71E-29	4.41E-146	0.00E+00
LIN-39	9.17E-01	1.32E-03	5.97E-13	2.34E-67	4.41E-05	1.12E-04	2.85E-58	7.04E-162
LIN-40	9.16E-01	1.79E-06	2.58E-03	4.16E-112	2.27E-03	7.43E-03	1.46E-39	9.42E-191
LIR-3	7.02E-01	1.70E-07	1.31E-09	1.19E-85	2.53E-04	9.01E-06	4.40E-70	3.24E-193
LSY-12	5.62E-01	1.58E-06	7.05E-03	8.75E-49	9.65E-03	9.40E-04	2.77E-30	1.80E-115
LSY-2	5.88E-01	3.08E-19	1.43E-04	3.40E-210	3.60E-03	1.68E-04	8.18E-145	0.00E+00
LSY-27	3.67E-01	2.81E-01	1.10E-06	4.35E-14	2.79E-03	7.58E-03	8.68E-12	3.35E-21

TF	Up 3	Up 6	Up 12	Up 30	Down 3	Down 6	Down 12	Down 30
M03D4.4	1.07E-01	8.03E-04	6.05E-02	1.51E-35	1.70E-01	6.17E-01	1.30E-25	1.89E-77
MAB-5	5.24E-02	1.32E-04	2.28E-07	1.19E-60	5.51E-01	1.96E-02	8.67E-46	1.28E-156
MADF-10	8.95E-01	1.15E-05	7.64E-04	1.61E-57	6.27E-04	7.83E-04	8.80E-33	9.65E-124
MADF-2	9.16E-01	1.02E-04	9.61E-03	1.42E-25	2.85E-03	6.28E-04	3.21E-19	1.87E-58
MBF-1	5.64E-01	7.70E-01	9.97E-05	1.03E-05	4.12E-01	5.57E-02	1.96E-04	1.61E-07
MDL-1	9.06E-01	2.89E-07	6.39E-01	1.92E-107	7.73E-02	4.16E-01	2.05E-36	8.39E-203
MEC-3	5.72E-01	6.89E-03	2.83E-05	1.54E-09	5.86E-03	2.32E-03	3.10E-15	2.08E-32
MED-1	3.74E-01	5.61E-03	1.12E-05	1.30E-16	1.42E-03	2.68E-01	7.84E-22	4.55E-51
MEF-2	6.31E-01	1.31E-02	4.02E-01	2.54E-21	2.14E-02	3.03E-02	4.88E-14	1.61E-48
MEL-28	4.90E-01	1.60E-03	1.24E-01	2.42E-07	4.99E-02	5.36E-02	1.77E-05	1.20E-17
MEP-1	6.37E-02	1.38E-14	1.48E-04	9.51E-185	2.98E-06	1.70E-07	6.80E-109	0.00E+00
MLS-2	2.93E-01	1.02E-07	3.42E-02	5.20E-42	1.09E-01	4.67E-03	2.77E-18	1.81E-89
MML-1	8.92E-01	5.83E-01	3.29E-02	9.51E-03	5.39E-01	8.91E-01	8.92E-01	1.93E-01
MXL-1	4.48E-01	3.50E-02	2.21E-02	8.92E-05	6.67E-01	1.00E+00	2.22E-03	9.98E-06
NFYA-1	3.52E-05	5.89E-09	1.37E-09	2.13E-171	3.14E-01	1.23E-02	7.95E-121	0.00E+00
NHR-10	1.65E-01	3.01E-06	9.17E-01	1.01E-27	2.16E-01	1.08E-01	4.40E-12	1.80E-43
NHR-102	5.54E-02	3.99E-05	7.85E-02	1.33E-54	5.20E-04	1.62E-01	1.43E-20	8.54E-90
NHR-11	9.51E-01	6.85E-05	3.83E-03	3.19E-58	4.39E-01	4.89E-01	6.34E-38	9.40E-125
NHR-116	8.30E-01	9.13E-01	6.28E-04	1.32E-07	1.39E-02	1.25E-01	1.06E-03	8.72E-12
NHR-12	8.35E-01	1.96E-03	2.89E-02	1.30E-15	2.08E-01	3.79E-03	4.45E-11	3.16E-35
NHR-129	3.51E-01	3.00E-15	9.31E-04	2.59E-142	4.70E-01	3.33E-01	3.05E-87	0.00E+00
NHR-179	8.19E-02	1.61E-02	3.24E-02	5.02E-02	1.00E+00	4.24E-01	6.45E-03	9.86E-02
NHR-2	8.92E-01	4.85E-04	6.01E-01	1.59E-37	1.79E-02	1.48E-03	1.39E-10	3.12E-73
NHR-20	6.58E-02	3.17E-15	7.45E-04	2.26E-146	6.91E-06	5.31E-04	2.84E-71	3.76E-293
NHR-21	3.73E-01	3.61E-04	1.06E-01	2.87E-49	2.21E-05	5.13E-03	3.47E-28	3.35E-127
NHR-23	3.70E-01	1.03E-08	3.75E-01	2.07E-72	8.94E-04	4.89E-04	2.72E-56	1.23E-244
NHR-232	4.37E-02	1.38E-02	6.83E-02	1.33E-12	1.56E-01	9.44E-01	2.01E-12	1.24E-26
NHR-237	6.71E-01	3.81E-12	4.60E-11	2.70E-75	3.20E-08	3.28E-15	1.65E-51	9.34E-183
NHR-25	2.99E-01	2.33E-07	1.32E-03	5.34E-64	4.22E-05	3.13E-05	5.69E-50	3.72E-194
NHR-28	1.94E-05	4.94E-12	2.13E-26	9.72E-61	1.10E-01	1.22E-06	7.62E-04	8.96E-87
NHR-43	9.77E-01	9.85E-04	5.64E-04	1.60E-48	8.12E-06	5.49E-02	2.10E-39	8.03E-129
NHR-47	1.81E-01	3.56E-11	8.95E-11	5.03E-143	7.03E-05	9.86E-06	2.20E-80	0.00E+00
NHR-48	8.41E-02	1.59E-01	4.73E-02	2.81E-26	7.63E-01	2.86E-02	4.91E-09	5.27E-44
NHR-6	3.02E-01	6.52E-16	6.70E-01	3.21E-89	6.25E-05	7.17E-09	1.28E-47	1.17E-218
NHR-67	4.94E-01	1.00E+00	4.65E-04	8.09E-08	9.21E-03	4.89E-02	5.32E-10	2.05E-17
NHR-71	7.64E-01	1.05E-08	3.55E-05	1.37E-138	5.27E-01	8.11E-01	3.08E-91	3.38E-294
NHR-76	4.56E-01	8.30E-04	1.63E-02	7.53E-29	9.36E-03	3.00E-05	1.66E-15	2.44E-46
NHR-77	7.58E-02	4.97E-12	6.94E-01	3.33E-78	1.16E-02	5.10E-01	4.78E-32	4.07E-152
NHR-80	1.19E-04	4.36E-15	1.51E-20	1.68E-84	2.05E-01	3.38E-08	7.92E-12	4.09E-130
NHR-84	1.72E-02	7.77E-01	7.68E-02	1.19E-03	1.53E-01	1.19E-01	3.38E-02	2.23E-05
NHR-85	3.77E-01	1.63E-05	6.46E-02	1.55E-62	3.00E-05	2.38E-06	4.03E-31	7.98E-124
NHR-90	9.59E-01	5.25E-08	7.03E-06	2.94E-104	6.88E-09	9.91E-07	3.39E-64	9.70E-244
NPAX-4	9.34E-01	3.49E-02	1.20E-03	7.81E-12	1.78E-02	1.76E-01	1.46E-15	6.41E-25
ODD-2	3.22E-01	6.89E-01	6.76E-03	6.29E-07	5.60E-02	1.08E-01	4.07E-10	9.30E-18
PAG-3	6.03E-02	2.67E-01	9.58E-03	4.63E-19	6.41E-01	2.28E-01	1.89E-25	4.26E-51
PAX-1	8.63E-01	9.02E-04	7.51E-07	2.15E-42	4.05E-07	6.25E-08	2.84E-28	1.57E-94
PEB-1	1.04E-02	1.19E-01	2.62E-02	6.86E-01	2.44E-02	4.74E-01	6.73E-01	5.40E-01
PHA-4	1.31E-01	1.81E-19	2.12E-01	3.33E-134	2.67E-01	1.81E-01	9.99E-74	0.00E+00
POP-1	6.57E-01	3.40E-02	1.14E-02	6.80E-17	9.33E-03	2.21E-01	6.79E-11	3.56E-39
PQM-1	3.69E-10	3.84E-11	3.54E-36	3.75E-15	7.26E-02	3.78E-12	8.60E-01	5.27E-15
R02D3.7	3.92E-01	5.02E-07	2.57E-14	1.44E-84	9.12E-11	1.69E-11	2.89E-69	1.05E-213
R06F6.6	6.82E-01	9.32E-02	1.54E-03	2.94E-02	7.24E-03	9.36E-02	1.05E-02	3.06E-03
RBR-2	7.11E-02	1.14E-07	6.22E-08	2.17E-107	1.65E-02	3.01E-07	7.66E-62	1.56E-243
REF-2	6.25E-01	1.69E-05	1.04E-02	2.07E-53	4.83E-03	1.11E-05	3.84E-36	7.79E-129
REPO-1	2.25E-01	2.66E-05	1.24E-03	6.09E-35	9.04E-02	5.21E-03	2.10E-20	1.61E-66

TF	Up 3	Up 6	Up 12	Up 30	Down 3	Down 6	Down 12	Down 30
RNT-1	1.17E-02	1.28E-07	3.08E-15	8.70E-130	6.00E-01	9.59E-02	6.46E-139	0.00E+00
SDC-2	4.98E-02	7.45E-03	4.72E-01	7.54E-33	3.73E-01	2.96E-01	3.90E-09	2.57E-48
SDZ-38	5.60E-01	3.25E-01	6.86E-01	4.03E-02	7.77E-02	1.60E-02	3.28E-01	2.60E-02
SEA-2	8.56E-01	6.63E-02	7.93E-04	3.44E-07	7.57E-01	9.69E-02	2.35E-01	2.38E-06
SEM-4	1.27E-01	1.04E-06	9.71E-05	6.65E-62	6.10E-02	2.45E-03	2.25E-35	1.66E-148
SKN-1	1.04E-01	9.31E-28	1.72E-03	5.12E-135	1.15E-08	8.20E-16	1.26E-78	0.00E+00
SMA-3	2.01E-01	3.58E-13	1.65E-02	2.45E-101	4.98E-12	1.58E-10	6.22E-56	3.73E-246
SMA-9	2.43E-02	3.03E-04	7.58E-02	2.02E-67	8.95E-01	1.01E-04	4.35E-37	3.70E-129
SNPC-4	3.88E-01	2.89E-19	6.99E-14	2.73E-149	1.83E-07	5.93E-13	4.50E-107	0.00E+00
SNU-23	2.29E-01	4.29E-10	9.20E-07	3.48E-175	1.05E-01	1.01E-01	5.33E-118	0.00E+00
SOX-4	9.41E-01	6.84E-05	1.35E-08	1.09E-54	1.70E-04	7.46E-05	2.87E-50	2.61E-127
SPR-1	7.09E-01	2.96E-03	2.64E-04	1.21E-42	9.77E-03	1.05E-03	8.48E-31	9.66E-87
SPTF-1	6.18E-02	6.00E-06	6.59E-05	1.22E-65	5.15E-03	5.97E-04	3.63E-33	2.52E-137
SWSN-7	1.42E-02	5.81E-14	4.22E-10	5.74E-156	3.15E-04	2.71E-06	8.71E-102	0.00E+00
SYD-9	8.13E-02	5.35E-02	8.74E-02	6.85E-20	6.10E-01	7.68E-02	4.10E-14	1.94E-36
T07F8.4	6.81E-01	4.89E-01	1.53E-03	4.93E-10	2.45E-01	5.32E-01	1.70E-12	1.24E-16
T26A5.8	7.08E-03	1.35E-03	2.84E-03	6.16E-09	4.82E-02	7.99E-03	4.22E-07	5.02E-13
TBX-2	4.09E-03	1.38E-07	3.52E-11	1.26E-150	4.23E-02	9.11E-03	6.53E-99	0.00E+00
TBX-7	7.26E-01	7.48E-05	2.49E-05	8.40E-36	8.43E-05	2.04E-06	2.49E-41	2.52E-92
TBX-9	2.96E-01	1.23E-02	2.21E-02	5.16E-08	2.30E-02	3.86E-01	3.46E-06	3.81E-13
TLP-1	3.13E-06	8.37E-03	1.52E-07	1.82E-01	1.75E-01	1.00E+00	5.64E-02	1.91E-01
TRA-4	5.27E-01	5.03E-03	3.07E-02	1.84E-10	2.92E-03	1.27E-02	1.42E-06	3.42E-21
TTX-3	9.55E-01	2.60E-02	6.01E-08	2.40E-26	4.82E-06	9.40E-05	7.42E-31	1.43E-65
UNC-120	8.07E-04	3.03E-03	7.11E-04	2.62E-65	3.85E-03	9.72E-03	2.65E-50	3.54E-145
UNC-130	3.66E-02	1.61E-09	2.08E-32	4.35E-152	3.88E-09	9.58E-13	4.27E-179	0.00E+00
UNC-3	2.46E-01	5.79E-03	1.56E-04	2.22E-16	1.23E-02	2.16E-02	6.44E-23	4.69E-50
UNC-30	5.43E-01	4.61E-01	9.69E-05	7.14E-09	1.66E-01	2.69E-01	4.73E-08	7.37E-17
UNC-39	5.38E-02	1.00E+00	7.92E-01	9.34E-10	5.06E-01	7.35E-02	2.64E-04	2.36E-16
UNC-42	4.32E-02	4.74E-08	2.22E-06	8.98E-90	9.50E-01	1.66E-04	3.21E-59	1.66E-220
UNC-55	1.69E-02	1.72E-17	5.07E-11	2.87E-167	2.43E-01	8.59E-04	2.05E-132	0.00E+00
UNC-62	7.94E-02	7.67E-04	2.98E-05	8.32E-29	3.08E-04	3.01E-11	5.15E-07	8.89E-37
UNC-86	1.47E-01	5.43E-07	1.03E-14	1.11E-114	5.65E-04	1.20E-08	3.02E-80	1.23E-274
VAB-15	5.47E-01	3.52E-03	1.90E-07	5.12E-39	4.49E-07	2.94E-07	6.25E-33	4.29E-97
W03F9.2	8.46E-01	2.00E-17	3.67E-20	3.38E-133	2.04E-19	5.67E-23	4.69E-109	0.00E+00
XBP-1	9.51E-03	3.54E-15	7.70E-14	8.29E-167	6.52E-09	5.57E-12	4.20E-116	0.00E+00
XND-1	7.02E-01	5.53E-05	7.12E-50	2.05E-109	4.77E-13	5.35E-27	2.83E-176	5.84E-275
Y116A8C.19	2.53E-01	2.39E-03	4.88E-05	9.56E-17	4.51E-05	1.37E-03	6.89E-20	4.27E-49
Y22D7AL.16	6.48E-01	5.42E-02	3.12E-03	2.33E-09	2.44E-02	2.80E-01	3.11E-10	7.28E-25
Y53C12C.1	8.24E-01	1.09E-06	5.95E-03	9.04E-27	9.62E-04	3.83E-03	5.08E-24	1.27E-67
ZAG-1	7.23E-01	2.58E-02	2.58E-02	1.82E-40	3.26E-01	4.02E-01	5.02E-18	3.40E-72
ZFP-2	5.16E-01	9.53E-02	1.26E-04	4.79E-31	1.99E-02	8.30E-03	4.52E-25	9.54E-55
ZIM-3	1.65E-01	8.88E-04	1.46E-03	1.36E-13	2.21E-02	1.79E-01	6.41E-11	1.00E-23
ZIP-2	1.74E-07	1.62E-11	7.83E-02	7.49E-18	1.38E-03	3.46E-04	5.78E-06	8.49E-33
ZIP-5	5.48E-01	5.72E-04	4.11E-12	2.66E-38	5.03E-06	4.53E-06	1.04E-51	4.35E-108
ZK185.1	3.88E-01	4.78E-04	3.34E-10	5.31E-73	1.66E-04	4.03E-05	1.18E-59	4.86E-145
ZK546.5	9.86E-02	2.31E-01	3.08E-09	1.06E-33	1.76E-04	1.74E-06	6.41E-26	3.12E-59
ZTF-11	7.31E-02	4.96E-05	2.15E-11	1.90E-89	4.37E-07	3.59E-09	6.38E-76	5.82E-216
ZTF-16	2.31E-01	1.32E-03	7.37E-04	5.66E-45	4.26E-03	1.92E-04	3.60E-30	3.05E-100
ZTF-18	6.37E-01	1.34E-06	2.28E-11	7.57E-68	3.12E-05	1.35E-08	8.03E-67	4.53E-165
ZTF-26	8.31E-01	5.12E-03	8.99E-02	4.61E-39	1.22E-01	5.19E-02	4.69E-11	4.89E-51
ZTF-4	6.56E-01	8.97E-05	1.22E-03	1.76E-62	6.19E-02	5.99E-03	1.97E-31	4.90E-121
ZTF-7	5.92E-01	7.35E-04	2.81E-01	1.62E-41	1.70E-02	4.42E-04	1.85E-18	1.40E-71

Chapter Four

Future Directions & Concluding Remarks

Future directions

The role of the intracellular pathogen response (IPR) in cellular reprogramming

A major finding from the analysis of the temporal transcriptional profile of transorganogenesis is that an IPR-like transcriptional response is strongly and transiently activated following ectopic expression of *elt-7* (Chapter 3; Figs 5, S10). Additionally, activation of the IPR is spatially restricted to the intestine, pharynx, and somatic gonad, the same starting and ending tissues involved in transorganogenesis. Ectopic expression of non-endoderm related transcription factors, i.e. *hlh-1* and *elt-1*, also activates the IPR in intestine, pharynx, and somatic gonad tissues, although no reprogramming is observed. Furthermore, the IPR appears to be activated specifically in response to the foreign transcriptional activity of these factors; ectopic expression of *hsp::nls::lacZ*, for instance, does not induce IPR activation (Tsunghan Yeh, personal communication). The strong correlation between the spatio-temporal activation of the IPR and somatic gonad-to-intestine transorganogenesis suggests a mechanistic link between these processes. However, a causative relationship has yet to be established.

Initial experiments have sought to test the hypothesis that activation of the IPR is necessary for ELT-7 directed reprogramming by inhibiting the IPR through RNAi of critical regulators. Unfortunately, the precise mechanism by which the IPR is activated or the upstream cellular stimulus to which it responds are poorly understood. Therefore, a small screen of factors known to be involved in the response to intracellular infection as well as innate immunity in general was attempted, but no significant effects on reprogramming efficiency were observed (Table 1). However, there are additional remaining factors that

were not tested for technical reasons. Also, RNAi inhibition of a rapidly and strongly activated gene may be insufficient for generating a loss of function phenotype; genetic mutants may yield more reliable results. Additionally, the IPR may involve factors with redundant or overlapping functions, so inhibition of multiple factors simultaneously may be necessary to alter any effects on reprogramming.

The activation of an innate immune response to intracellular pathogen infection suggests an intriguing mechanism for promoting direct cellular reprogramming. In order for a differentiated cell to acquire a disparate fate and undergo dramatic structural reorganization and remodeling, the preexisting proteins need to be removed and recycled. Defense against microsporidia involves ubiquitylation components, the proteasome, and the autophagy pathway (Bakowski et al., 2014). For example, one of the critical IPR components in *N. parisii* immunity is *cul-6*, a member of an SCF multi-subunit E3 ubiquitin ligase complex.

In addition to the IPR overall, proteasome components were found to be significantly upregulated following ectopic *elt-7* expression. A broad array of Reactome pathway terms are enriched among upregulated DEGs 6 hrs PHS and visualizing their interaction network reveals a common set of proteasome-associated factors (Fig 1). The gene expression paths for these factors shows that the proteasome is dynamically regulated during transorganogenesis, with rapid induction initiated after 3 hrs PHS and a peak in expression at 6 hrs PHS (Fig 2). These proteasomal genes are also upregulated in controls at 12 and 30 hrs PHS, most likely owing to their role in embryogenesis, which includes the regulation of developmental factors that determine proper cell fate acquisition (Du et al., 2015).

The ubiquitin-proteasome system and autophagy are highly conserved and critical for maintenance of cellular homeostasis (Dikic, 2016; Pohl and Dikic, 2019). These processes

form an interconnected quality control network that can function in a cell autonomous and nonautonomous manner in response to cellular stress (Poillet-Perez and White, 2019). If the IPR is critical for transorganogenesis, it will be interesting to find out the nature of the role it plays. Does the intestine gene regulatory network hijack the protein degradation pathways activated by cellular stress to promote the establishment of intestine cell fate? Or, does ectopic transcription factor over-expression cause a type of "autoimmune" response that destabilizes the endogenous cell state allowing it to be commandeered by the more robust intestine GRN? If the former hold true, then activating the IPR pathways in additional tissues may allow for reprogramming of other cell types into intestine. In the latter scenario, expression of additional reprogramming factors may be sufficient for converting the pharynx and somatic gonad into tissue other than intestine.

Analysis of transcript-level RNA-seq data

The transorganogenesis temporal transcriptional profile analysis (Chapter 3) was performed on gene-level transcript abundances, i.e. the overall expression level of the *aap-1* gene. The sequencing data from each RNA library, however, contains information on the abundance of each mRNA isoform in the *C. elegans* transcriptome, i.e. the expression levels for the Y110A7A.10.1 and Y110A7A.10.2 transcripts of the *aap-1* gene. Gene-level analysis was performed in part because of the more universal recognition of gene identifiers across databases. Many of the bioinformatic analysis tools used in this study are currently unable to process transcript-level data.

The additional dimension to the RNA-seq data possess a virtually untapped level of detail that has yet to be analyzed and could provide novel insight into the process of somatic gonad-

to-intestine transorganogenesis and other cellular consequences of ectopic ELT-7. Differential regulation of mRNA processing has been implicated in tissue specification in *C. elegans*, and some protein isoforms have tissue-specific expression patterns (Zahler, 2012). ELT-7 directed cell fate switching may therefore be facilitated by alternative splicing of mRNA isoforms, particularly if there are pharynx or uterus expressed genes that also have an intestine-specific isoform. Differential splicing has also been found to be involved in the regulation of human reprogramming of fibroblasts into induced pluripotent cells; the SFRS11 gene regulates splicing of genes that are critical for reprogramming (Toh et al., 2016).

Initial analysis of RNA-seq data at the 3 hrs PHS time point shows that, indeed, there is differential expression of individual isoforms of the same gene. There are approximately 150 genes at 3 hrs PHS that have at least one upregulated transcript and one downregulated transcript. Also, interestingly, mRNA splicing related pathways are enriched among differentially expressed genes upregulated 6 hrs PHS (Fig 1). It is possible that during early reprogramming ectopic ELT-7 activates the transcription of intestine-specific splicing factors, leading to a second phase in which endogenous RNAs are differentially spliced to produce intestine proteins, thereby accelerating the reprogramming process within tissues expressing a susceptible mRNA profile.

Unexplored relationships with significantly correlated data sets

The correlation between ELT-7 driven cellular reprogramming and the activation of the intracellular pathogen response (IPR) was discovered by searching a database of *C. elegans* gene expression studies (Yang et al., 2016) that had significant overlap with DEGs whose variance had the greatest contribution to the first principle component in the overall

transcriptional profiling analysis (Chapter 3; Figs 1E, 5, S10). Although genes differentially expressed to due intracellular pathogen infection had the most significant correlation to ELT-7 reprogramming-related DEGs, there are many additional studies and data sets that are also very strongly correlated (Table 2).

One group of data with a high degree of similarity to ELT-7 driven transcriptional changes are genes that are significantly upregulated during normal aging, in Day 3, 5 and 10 adults (p-values $\leq 5.30 \times 10^{-35}$) (Rangaraju et al., 2015). The aging transcriptome study found that as worms age there is increasing variance in gene expression, which they refer to as "transcriptional drift." Furthermore, this effect can be attenuated through treatment with the antidepressant drug mianserin, which inhibits serotonergic signaling, resulting in increased lifespan (Rangaraju et al., 2015). Out of the 221 ELT-7 reprogramming associated DEGs used for comparison, 122 are associated with aging. Many of these genes belong to *fbxa*/F-box and *pals* classes that, interestingly, are also upregulated as part of the IPR.

Under normal conditions, IPR genes are not expressed, and worms containing the *pals-5p::gfp* IPR reporter transgene do not show fluorescence noticeably higher than background. However, in progressively older adult worms, some *pals-5p::gfp* expression is occasionally observed in individual intestinal cells (Chapter 3; Fig 5, and data not shown). In addition to defense against microsporidia and virus infection, IPR genes are also involved in proteostasis and thermotolerance (Reddy et al., 2017). It may be worthwhile to investigate the association of reprogramming-related gene expression changes with aging-related transcriptional drift and IPR-mediated proteostasis. For instance, in addition to cellular reprogramming, ectopic *elt-7* expression also results in larval arrest and lethality within ~3-4 days; does ELT-7 cause increased transcriptional drift and therefore accelerated aging? Is this an important feature of

direct cellular reprogramming? If so, then would mianserin treatment delay reprogramming-induced lethality, or inhibit the reprogramming process?

The temporal transcriptional profile of aging is just one of many other statistically significantly related studies (Table 2). Differential gene expression due to mutations in transcriptional regulators, chromatin factors, and other microbial infections also overlaps with ectopic ELT-7 induced DEGs. Further exploration of these relationships may provide deeper insight into the regulation of transorganogenesis.

Continuation of the reverse genetic screen

The RNAi screening results described in Chapter 2 focused on the most striking phenotypes, specifically, the inhibition of transorganogenesis from knockdown of the *epc-1, smo-1,* and *pyp-1* genes. Additional screening results, however, may merit follow-up experiments as well. Although less dramatic, a modest but reproducible reduction in reprogramming efficiency could also indicate that a factor is important in the reprogramming process. At least 10 other RNAi results fall into this category, including the potentially promising hit *dnj-11,* which encodes a DNAJ domain containing protein orthologous to the mammalian ZRF1/MIDA1/ MPP11/DNAJC2 family of ribosome-associated molecular chaperones. *dnj-11* functions in asymmetric cell division, regulation of cell fate determination, and response to metallic stress (Hatzold and Conradt, 2008; Sahu et al., 2013; Yan et al., 2013) and could be mediating the transcriptional response to ectopic ELT-7.

Over 220 chromatin related genes were screened for effects on reprogramming efficiency (Chapter 2; Fig 7). However, this only represents a subset of all potential chromatin factors – the *C. elegans* genome may contain ~800 genes that are associated with chromatin

(Hajduskova et al., 2018). The RNAi screens performed in this work involved scoring of fine anatomical detail and manual categorization of often subtle reprogramming phenotypes, thereby limiting the overall scale. The remaining chromatin factors could be more rapidly tested using a screening approach with a more straightforward phenotypic readout, such as suppression of larval arrest after ectopic *elt-7* expression. However, this type of approach would likely be less sensitive in the detection of true positive hits. The detailed screening methods in this work yielded an ~1% hit rate, so a simpler screen may find relatively few additional factors.

Alternatively, a more efficient, sensitive screening approach would involve automated quantification of ectopic intestine formation based on IFB-2::CFP fluorescence intensity. This is not feasible in the current system because of the high level of IFB-2::CFP expression in the endogenous intestine that can easily mask the ectopic intestine tissue signal from automated fluorescence detection. Although, more sophisticated genetic techniques could be employed to remove endogenous *ifb-2::cfp* expression. For example, generation of a conditional knockout of the *ifb-2::cfp* construct in early intestine development while leaving the construct intact in all other tissues would result in CFP expression exclusively in tissues that undergo transdifferentiation. Such a system would be far more amenable to automation and could potentially be utilized for genome wide RNAi screening for factors that regulate transorganogenesis.

Concluding remarks

What determines reprogramming ability – the pioneers or the topology?

The ability for ELT-7 to directly reprogram the pharynx and somatic gonad into intestine appears to depend upon a cellular context in which (a) the PHA-4/FOXA transcription factor has been active during development (Riddle et al., 2016) and (b) the tissue is capable of initiating the IPR (Chapter 3). But what could be the molecular basis underlying these contextual features? Transcription factors that exhibit the highest reprogramming activity often have the special ability to bind target sites on nucleosomal DNA, thus behaving as "pioneer factors" to activate genes within closed chromatin (Iwafuchi-Doi and Zaret, 2014). Successful cellular reprogramming also often involves the manipulation of multiple factors, at least one of which possesses pioneer factor ability, or the cell may already express a critical pioneer transcription factor (Zaret and Mango, 2016). Generation of induced pluripotent stem cells (iPSCs) from differentiated fibroblasts, for example, relies heavily on pioneer factors; Oct4, Sox2, and Klf4 all bind as pioneer factors to closed chromatin in early stages of iPSC reprogramming (Soufi et al., 2012).

Pioneer factors are also critical during development and a particularly well studied example in humans are the FOXA2 and GATA4 transcription factors, which bind to and open chromatin during liver development and are the initial activators of liver-specific genes (Cirillo et al., 2002). The cell type-specific activity of pioneer factors and local chromatin effects have been described recently in molecular detail (Donaghey et al., 2018). In a ChIP-seq assay, FOXA2 occupancy is restricted to a subset of motif-containing loci depending on the cell type it is expressed in, but it weakly interacts, or "samples," potential binding regions

broadly. Sampling of target regions by FOXA2 can be stabilized by co-expression with GATA4, suggesting cooperative binding as a strong influence on cell type-specific DNA binding. However, FOXA2 and/or GATA4 binding does not correlate with increased chromatin accessibility, and therefore pioneer transcription factor occupancy alone may do little to modify chromatin structure (Donaghey et al., 2018).

Co-expression of critical pioneer transcription factors is an attractive theoretical mechanism for determining reprogramming ability. Intriguingly, PHA-4 and ELT-7 are homologs of FOXA2 and GATA4, respectively, and the combined endogenous PHA-4 and ectopic ELT-7 expression in the pharynx and somatic gonad might enable reprogramming into intestine. However, widespread ectopic expression of both PHA-4 and ELT-7 was not sufficient for reprogramming of additional tissues (Riddle et al., 2016). There may be other endogenously expressed factors that co-operate with ELT-7 to drive transorganogenesis, although RNAi screening of all Forkhead and GATA factors did not reveal any such interactions (Chapter 2). While there may still be important reprogramming cofactors expressed in the pharynx and somatic gonad, pioneer factors alone are likely insufficient for reprogramming in all contexts.

Within the nucleus, chromatin has a hierarchically organized architecture that may be a primary determinant of possible reprogramming trajectories (Guo and Morris, 2017). Sub-chromosomal regions are compartmentalized, and co-regulated elements form topologically associated domains (TADs) (Dixon et al., 2012). Some highly compacted regions and/or compartmentalized TADs may be inaccessible even to pioneer factors. For example, in early iPSC reprogramming, there are megabase scale H3K9me3-containing genomic regions that are refractory to binding by the OSKM factors (Soufi et al., 2012).

The ability of ELT-7 to reprogram the pharynx and somatic gonad into intestine may be attributable to topological similarities among these particular tissues. Although distantly related in cell lineage, convergent developmental events could lead to nuclear architectures that bare enough resemblance to one another to permit cell fate switching between pharynx/uterus and intestine. As nuclear organization is dynamic during development, dependence on chromatin topology could explain why pharynx-to-intestine transdifferentiation is dependent on differentiated pharyngeal tissue identity, and possibly account for the restricted developmental window in which somatic gonad-to-intestine transorganogenesis is permissible. For example, in immature pharyngeal tissue there may be chromatin compartments that ELT-7 is unable to access that become less restricted by terminal differentiation. Or, in the developing somatic gonad, perhaps there are intestine-like TADs that temporarily form that can be acted upon by ELT-7 to induce cell fate change, but then later become dissociated or sequestered.

Direct cellular reprogramming may also require considerable restructuring of nuclear architecture. In transorganogenesis, this could be driven by the endoderm GATA factors themselves. In addition to DNA binding transcription factor activity, GATA factors have been shown to be capable of facilitating chromosome looping and mediating long-range control of gene expression (Vakoc et al., 2005). It has been proposed that the different modes of GATA factor activity are concentration dependent, with local, single site binding at low protein concentrations and multi-site, long-range binding at high concentration (Chen et al., 2012). Direct reprogramming to intestine by ELT-7 involves high levels of ectopic expression – is ELT-7 capable of remodeling chromatin topology at high concentration? What is the dosage dependency of transorganogenesis? Chromatin reorganizing abilities of

GATA factors has not been demonstrated in *C. elegans*, so it remains to be seen whether this mode of reprogramming is possible. Ultimately, both pioneer transcription factors and nuclear architecture are likely contributing to the ability of ELT-7 to direct cell fate changes.

Towards complete control of cell fate

The conversion of one organ into another within a living animal, initiated by a single genetic switch, is an awe-inspiring phenomenon. It is not simply cells cultured *in vitro* which are slowly coaxed into adopting a different cell fate, but a rapid and thorough restructuring *in vivo* with near deterministic efficiency. How exactly this is possible in *C. elegans* is yet to be fully understood, and will require further research into the fundamental mechanisms that control cellular identity. For instance, is the maintenance of cell fate an active, continuous process in all differentiated tissues, or do different tissues employ different strategies? Also, can a transcription factor that positively directs acquisition of a particular cell fate act simultaneously to repress alternative cell fates? The work presented in Chapters 2 and 3 sheds some light on these questions.

When *elt-7* is expressed ectopically throughout the worm, the transcriptional effects vary dramatically depending upon the cell type. At one extreme, in the pharynx and somatic gonad, the cells are completely reprogrammed into intestine-like tissue. Conversely, in some tissues, such as A-class neurons, the average transcriptional state is virtually unperturbed. Interestingly, there appears to be a range of responses in between, with many tissues that are strongly impacted by ectopically expressed *elt-7* but able to maintain their established fates (Chapter 3; Figs 3, S3). This finding reveals that in regard to cellular reprogramming by

ectopically expressed factors, even though cells may not display any phenotypic changes, there could still be a significant impact on their transcriptional circuitry.

The tissue-specific variability to ectopic *elt-7* expression suggests that different mechanisms are utilized in the maintenance of cell fate, and therefore distinct cell types possess a unique set of barriers to forced reprogramming (Fig 3). Pharynx cells are the least resistant to ELT-7 and are rapidly reprogrammed. The somatic gonad is permissible to reprogramming during a restricted temporal window and requires some additional time to undergo transorganogenesis (Chapter 2; Fig 2). Ectopic intestine marker expression is occasionally observed at low levels in other tissues that may also have transiently repressed transcription. The ability of ELT-7 to reprogram additional tissues into intestine is therefore a likely possibility. Further research into ELT-7 reprogramming and transorganogenesis will hopefully identify the critical barriers, and we may eventually be able to convert the entire worm into a gut.

Table 1. Preliminary innate immunity RNAi results.

Gene	RNAi confirmed	Description	Rationale	Reference	Effect on Repro.
pmk-1	Yes	p38 Map Kinase family	initiator of innate immune response	Millet and Ewbank 2004	no change
sek-1	*low qual*	SAPK/ERK kinase, MAPKK	initiator of innate immune response	Millet and Ewbank 2004	no change
nsy-1	tsn-1 (F10G7.2)	Nuronal SYmmetry, MAPKKK	initiator of innate immune response	Millet and Ewbank 2004	*n.d.*
dbl-1	Yes	DPP/BMP-Like, TGFbeta homolog	initiator of innate immune response	Millet and Ewbank 2004	no change
zip-2	no hit?	bZIP transcription factor	required for P. aeruginosa immune response	Ewbank and Pujol 2016	*n.d.*
atfs-1	nhr-42	Activating Transcription Factor associated with Stress, bZIP factor	required for mitochondrial UPR	Ewbank and Pujol 2016	*n.d.*
cul-6	Yes	Cullin, component of SCF E3 Ub ligase complex	intracellular pathogen response gene	Bakowski et al. 2014	no change
skr-3	Yes, and skr-4	Skp1 related, Ub ligase complex component	intracellular pathogen response gene	Bakowski et al. 2014	no change
skr-4	*no clone*	Skp1 related, Ub ligase complex component	intracellular pathogen response gene	Bakowski et al. 2014	*n.d.*
skr-5	Yes	Skp1 related, Ub ligase complex component	intracellular pathogen response gene	Bakowski et al. 2014	no change
let-363	*no clone*	mTOR homolog	let-363 RNAi increases autophagy	Bakowski et al. 2014	*n.d.*
pals-5	Yes	Protein containing ALS2CR12 signature	intracellular pathogen response gene	Bakowski et al. 2014	no change
F56A8.3	Yes	leucine rich repeat containing protein	F56A8.3 RNAi reduces larval arrest from intracellular infection	Luallen et al. 2015	unclear/no change
pals-22	Yes	Protein containing ALS2CR12 signature	inhibits expression of many IPR genes	Reddy et al. 2017	no change
mdl-1	Yes	MAD-Like, bHLH protein, Myc interaction network	transcriptional overlap with PHA-4, inhibition reduces intracellular pathogen load	Botts et al. 2016	no change
mxl-1	Yes	MaX-Like, bHLH protein, Myc interaction network	can heterodimerize with MDL-1, bind E-box DNA sequence	Botts et al. 2016	no change
mxl-2	Yes	MaX-Like, bHLH protein, Myc interaction network	inhibition reduces intracellular pathogen load	Botts et al. 2016	no change
mml-1	T07E3.4	Myc and Mondo-Like, bHLH-ZIP protein	inhibition increases intracellular pathogen load	Botts et al. 2016	*n.d.*
zip-10	Yes	bZIP transcription factor	IPR upregulated gene, may act in pathway with PHA-4/MDL-1	Botts et al. 2016	no change

n.d., not determined.

Fig 1.

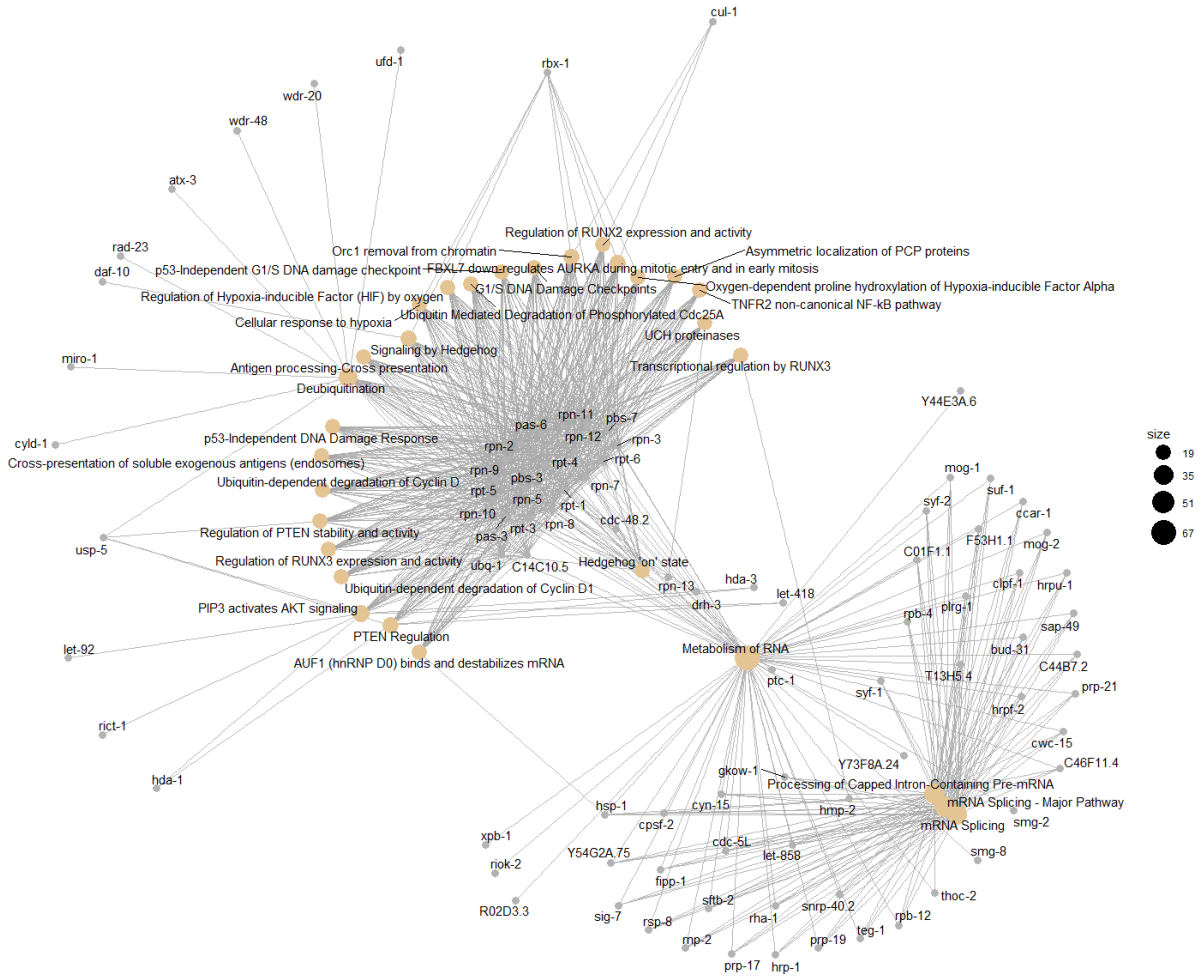

cul-1
ufd-1
rbx-1
wdr-20
wdr-48
atx-3
rad-23
daf-10
Regulation of RUNX2 expression and activity
Orc1 removal from chromatin
Asymmetric localization of PCP proteins
p53-Independent G1/S DNA damage checkpoint
FBXL7 down-regulates AURKA during mitotic entry and in early mitosis
G1/S DNA Damage Checkpoints
Oxygen-dependent proline hydroxylation of Hypoxia-inducible Factor Alpha
Regulation of Hypoxia-inducible Factor (HIF) by oxygen
TNFR2 non-canonical NF-kB pathway
Ubiquitin Mediated Degradation of Phosphorylated Cdc25A
Cellular response to hypoxia
UCH proteinases
Signaling by Hedgehog
miro-1
Antigen processing-Cross presentation
Transcriptional regulation by RUNX3
Deubiquitination
Y44E3A.6
rpn-11
pbs-7
pas-6
rpn-12
rpn-3
cyld-1
p53-Independent DNA Damage Response
rpn-2
rpt-4
rpt-6
Cross-presentation of soluble exogenous antigens (endosomes)
rpn-9
pbs-3
Ubiquitin-dependent degradation of Cyclin D
rpt-5
rpn-7
rpn-5
rpt-1
rpn-10
cdc-48.2
rpt-3
rpn-8
pas-3
Regulation of PTEN stability and activity
usp-5
Regulation of RUNX3 expression and activity
ubq-1
C14C10.5
Hedgehog 'on' state
rpn-13
hda-3
let-418
Ubiquitin-dependent degradation of Cyclin D1
drh-3
PIP3 activates AKT signaling
PTEN Regulation
let-92
AUF1 (hnRNP D0) binds and destabilizes mRNA
Metabolism of RNA
rict-1
hda-1
ptc-1
syf-1
mog-1
suf-1
syf-2
ccar-1
F53H1.1
mog-2
C01F1.1
clpf-1
hrpu-1
plrg-1
rpb-4
sap-49
bud-31
C44B7.2
T13H5.4
prp-21
hrpf-2
cwc-15
C46F11.4
Y73F8A.24
gkow-1
Processing of Capped Intron-Containing Pre-mRNA
mRNA Splicing - Major Pathway
hsp-1
cyn-15
hmp-2
cpsf-2
mRNA Splicing
smg-2
xpb-1
cdc-5L
smg-8
Y54G2A.75
let-858
riok-2
fipp-1
thoc-2
sftb-2
snrp-40.2
R02D3.3
sig-7
rsp-8
rha-1
rpb-12
teg-1
rnp-2
prp-19
prp-17
hrp-1
size
19
35
51
67

Fig 1. Enriched pathway ontologies among upregulated DEGs 6 hrs PHS.

Interaction network between differentially expressed genes (grey nodes) and their associated biological pathways (orange nodes) in the Reactome database. Pathway node size is proportional to the number of connected DEGs. A diverse set of pathways are related through a common set of upregulated proteasomal genes. Also upregulated at 6 hrs PHS are numerous mRNA splicing-related genes. For simplicity, only the top 30 most significant pathway terms are shown out of 89 enriched terms. See also Chapter 3, Fig S6.

Fig 2.

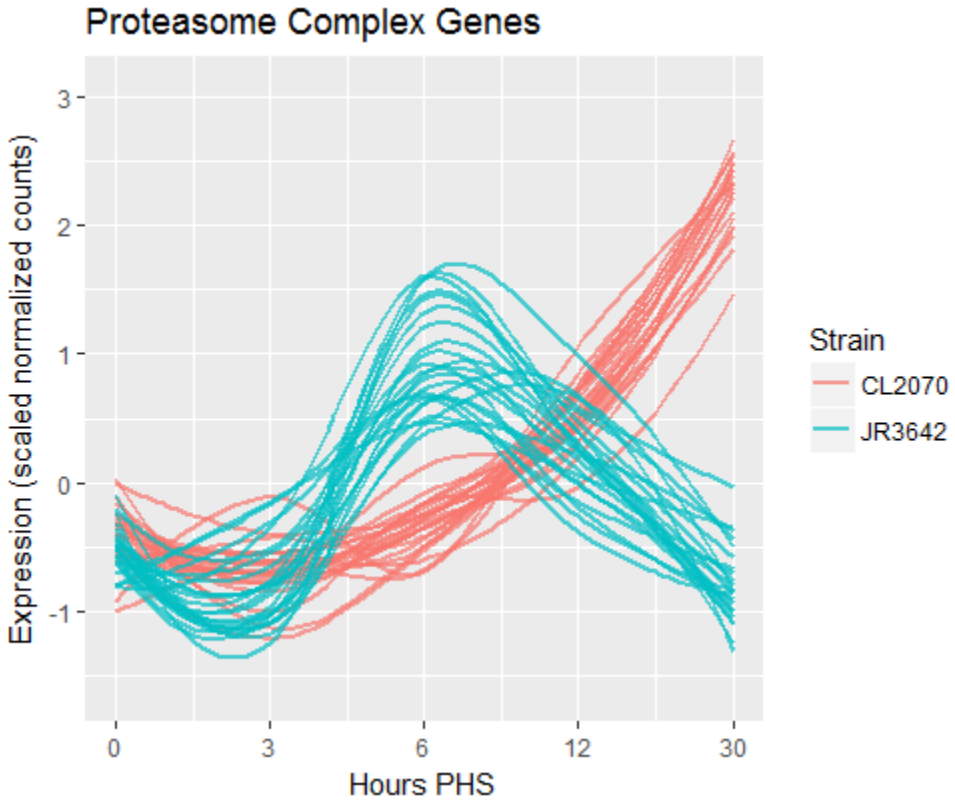

Proteasome Complex Genes
Expression (scaled normalized counts)
3
2
1
0
-1
0
3
6
12
30
Hours PHS
Strain
CL2070
JR3642

Fig 2. Proteasomal genes are upregulated during transorganogenesis.

Gene expression paths for factors associated with the "proteasome complex" cellular component ontology term, which is enriched among upregulated DEGs at 6 hrs PHS. Upregulation of proteasomal activity may enhance protein turnover, which could facilitate cell fate change and transorganogenesis. See also Chapter 3, Fig S4. CL2070, *hsp::gfp* control; JR3642, *hsp::elt-7*.

Table 2. Top 50 data sets most significantly correlated with high variance genes post ectopic ELT-7 expression.

Category	Term	Counts	ListSize	PopHit	Pop Size	Pvalue	PMID
Microbes	UP by N. parisii at 8h	61	221	121	19857	1.14E-86	24945527
Microbes	UP by N. parisii at 16h	51	221	67	19857	1.60E-85	24945527
Microbes	UP by N. parisii at 30h	53	221	101	19857	5.18E-76	24945527
Microbes	UP by virus Orsay	32	221	49	19857	3.03E-49	23811144
Mutants	UP lin-15b and lin-35 mutant	55	221	320	19857	2.28E-48	21343362
Development/ Dauer/Aging	UP by aging on day 3	103	221	2001	19857	1.28E-43	26623667
Development/ Dauer/Aging	UP by aging on day 5	111	221	2451	19857	2.13E-42	26623667
Mutants	UP hpl-2 KO	70	221	817	19857	9.83E-42	22083954
Microbes	UP by N. parisii at 40h	30	221	64	19857	3.33E-40	24945527
Mutants	UP by dpy-21 mutant L3	46	221	288	19857	1.30E-38	26641248
Mutants	UP by elt-2 RNAi L4	31	221	89	19857	8.32E-37	27070429
Development/ Dauer/Aging	UP by aging on day 10	112	221	2988	19857	5.30E-35	26623667
Microbes	UP by N. parisii, array	28	221	78	19857	1.44E-33	24945527
Mutants	UP L4 hpl-2;hpl-1 and his-24 vs. wildtype	48	221	427	19857	3.59E-33	23028351
Mutants	UP by lin-35 mutant (Petrella)	55	221	619	19857	5.70E-33	21343362
Mutants	UP lin-15b mutant	72	221	1250	19857	9.19E-32	21343362
Other	affected by copy number variants	65	221	1138	19857	3.65E-28	20100350
DAF/Insulin/ food	Down nhr-114 RNAi vs. glp-1 mutant	53	221	913	19857	6.27E-23	23499532
Development/ Dauer/Aging	UP in day 12 vs. day 2 (Pu)	58	221	1161	19857	4.98E-22	25838541
Chemicals/ stress	UP by cadmium	29	221	227	19857	2.79E-21	17592649
Development/ Dauer/Aging	Down lin-35; mes-4 RNAi	29	221	231	19857	4.52E-21	21343362
Mutants	UP let-418 RNAi	44	221	666	19857	5.30E-21	21060680
Mutants	down by mec-3(e1338) mutant	31	221	335	19857	1.02E-18	23889932
Mutants	UP glp-1 mutant	53	221	1176	19857	4.50E-18	23499532
Microbes	UP by B. thuringiensis at 6h (BT247, 1:10) (Yang)	42	221	724	19857	5.56E-18	25720978
Microbes	UP by B. thuringiensis at 6h (BT247, 1:2) (Yang)	39	221	620	19857	7.78E-18	25720978
Mutants	down by mec-3(e1338), PVD/OLL	32	221	397	19857	1.33E-17	23889932
Development/ Dauer/Aging	UP by bar-1(ga80), with development effect	31	221	392	19857	8.05E-17	24819947
Other	Cluster 11, decreasing expression postembryonic	42	221	820	19857	4.35E-16	24036951

Category	Term	Counts	ListSize	PopHit	Pop Size	Pvalue	PMID
Mutants	UP by lin-35 mutant (Grishok)	33	221	513	19857	2.45E-15	19073934
DAF/Insulin/ food	down by rapamycin under eat-2 mutant	25	221	287	19857	1.69E-14	26676933
Tissue	Positive Enriched Intestinal genes	67	221	2358	19857	4.98E-13	22467213
Mutants	UP by aak-2 overexpressed and daf-2:rsks-1 mutant	42	221	1013	19857	5.15E-13	26959186
Other	UP by male in hermaphrodites	24	221	306	19857	5.81E-13	24292626
Mutants	down by wrn-1 mutant	39	221	899	19857	1.13E-12	25346348
Mutants	Down xpa-1 mutant	41	221	1011	19857	2.12E-12	20382984
Chemicals/ stress	UP by UV irradiation on xpa-1 mutant	37	221	889	19857	1.64E-11	25419847
Microbes	UP by N. parisii at 64h	13	221	79	19857	7.12E-11	24945527
TF Targets	HPL-2 L1 targets (Araya)	15	221	123	19857	9.80E-11	25164749
Mutants	UP lin-35 mutant(n745, NULL) L1	41	221	1154	19857	1.22E-10	17368442
Chemicals/ stress	Down by Beta-naphthoflavone 8	62	221	2415	19857	3.70E-10	11697852
Mutants	UP by xrn-2 RNAi	32	221	759	19857	4.02E-10	27631780
Chemicals/ stress	Cadmium-responsive	13	221	92	19857	4.44E-10	17592649
Microbes	UP by B. thuringiensis at 12h (BT247, 1:2) (Yang)	37	221	1010	19857	5.62E-10	25720978
Mutants	UP lin-35 mutant(n745, NULL) L4	41	221	1237	19857	9.40E-10	17368442
Mutants	UP in lin-35 mutant and down in lin-35;mes-4	11	221	60	19857	1.10E-09	21343362
Mutants	UP daf-7 mutant	21	221	336	19857	1.27E-09	15380030
Mutants	UP daf-14 mutant	26	221	556	19857	3.41E-09	15380030
Microbes	UP by B. thuringiensis at 12h (BT247, 1:10) (Yang)	42	221	1349	19857	3.43E-09	25720978
Mutants	UP L3 hpl-2 mutant	24	221	503	19857	1.10E-08	22185090

Pop Size, background complete gene list; PopHit, number of genes associated with the Term; ListSize, number of ELT-7 reprogramming DEGs used as input; Counts, number of genes associated with the Term and input DEGs; PMID, PubMed identifier for reference study.

Fig 3.

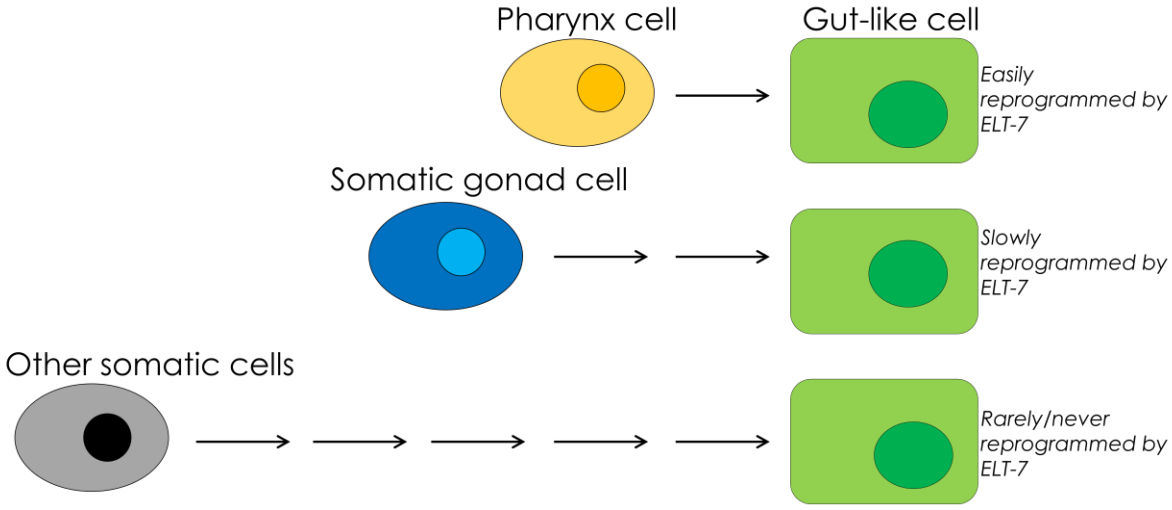

Pharynx cell
Gut-like cell
Easily reprogrammed by ELT-7
Somatic gonad cell
Slowly reprogrammed by ELT-7
Other somatic cells
Rarely/never reprogrammed by ELT-7

Fig 3. Conceptual model for cellular specificity of ELT-7 directed reprogramming. Pharynx and somatic gonad cells are effectively reprogrammed into gut-like cells following ectopic *elt*-7 expression. All other somatic cell types are refractory to reprogramming into intestine with some rare, weak intestine marker expression. Arrows represent unknown stochastic barriers that must be overcome in order for reprogramming to be permanent. The switch from one cell type to another may be 'easier' depending on the initial cellular context.

References

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B. A., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* **46**, 379.

Ahringer, J. and Gasser, S. M. (2018). Repressive Chromatin in Caenorhabditis elegans: Establishment, Composition, and Function. *Genetics* **208**, 491–511.

An, J. and Blackwell, K. T. (2003). SKN-1 links C. elegans mesendodermal specification to a conserved oxidative stress response. *Gene Dev* **17**, 1882–1893.

Anokye-Danso, F., Anyanful, A., Sakube, Y. and Kagawa, H. (2008). Transcription factors GATA/ELT-2 and forkhead/HNF-3/PHA-4 regulate the tropomyosin gene expression in the pharynx and intestine of Caenorhabditis elegans. *Journal of molecular biology* **379**, 201–11.

Auer, P. L. and Doerge, R. (2010). Statistical Design and Analysis of RNA Sequencing Data. *Genetics* **185**, 405–416.

Bakowski, M. A., Desjardins, C. A., Smelkinson, M. G., Dunbar, T. L., Dunbar, T. A., Lopez-Moyado, I. F., Rifkin, S. A., Cuomo, C. A. and Troemel, E. R. (2014). Ubiquitin-mediated response to microsporidia and virus infection in C. elegans. *PLoS pathogens* **10**, e1004200.

Baroux, C., Autran, D., Gillmor, C. S., Grimanelli, D. and Grossniklaus, U. (2008). The Maternal to Zygotic Transition in Animals and Plants. *Cold Spring Harb Sym* **73**, 89–100.

Baugh, L., Hill, A. A., Slonim, D. K., Brown, E. L. and Hunter, C. P. (2003). Composition and dynamics of the Caenorhabditis elegans early embryonic transcriptome. *Development (Cambridge, England)* **130**, 889–900.

Beagan, J. A., Gilgenast, T. G., Kim, J., Plona, Z., Norton, H. K., Hu, G., Hsu, S. C., Shields, E. J., Lyu, X., Apostolou, E., et al. (2016). Local Genome Topology Can Exhibit an Incompletely Rewired 3D-Folding State during Somatic Cell Reprogramming. *Cell Stem Cell* **18**, 611–624.

Becker, J. S., Nicetto, D. and Zaret, K. S. (2016). H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends Genet* **32**, 29–41.

Becker, S. F. and Jarriault, S. (2016). Current Opinion in Genetics & Development. *Cell reprogramming, regeneration and repair, edited by Peter Reddien and Elly Tanaka* **40**, 154–163.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J Royal Statistical Soc Ser B Methodol* **57**, 289–300.

Betschinger, J., Mechtler, K. and Knoblich, J. A. (2006). Asymmetric Segregation of the Tumor Suppressor Brat Regulates Self-Renewal in Drosophila Neural Stem Cells. *Cell* **124**, 1241–1253.

Blankenberg, D., Kuster, G., Bouvier, E., Baker, D., Afgan, E., Stoler, N., Team, G., Taylor, J. and Nekrutenko, A. (2014). Dissemination of scientific software with Galaxy ToolShed. *Genome Biol* **15**, 403.

Boeck, M. E., Boyle, T., Bao, Z., Murray, J., Mericle, B. and Waterston, R. (2011). Specific roles for the GATA transcription factors end-1 and end-3 during C. elegans E-lineage development. *Developmental biology* **358**, 345–55.

Bokman, S. H. and Ward, W. W. (1981). Renaturation of Aequorea green-fluorescent protein. *Biochem Bioph Res Co* **101**, 1372–1380.

Bolger, A. M., Lohse, M. and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120.

Borkent, M., Bennett, B. D., Lackford, B., Bar-Nur, O., Brumbaugh, J., Wang, L., Du, Y., Fargo, D. C., Apostolou, E., Cheloufi, S., et al. (2016). A Serial shRNA Screen for Roadblocks to Reprogramming Identifies the Protein Modifier SUMO2. *Stem Cell Reports* **6**, 704–716.

Bossinger, O., Fukushige, T., Claeys, M., Borgonie, G. and McGhee, J. D. (2004). The apical disposition of the Caenorhabditis elegans intestinal terminal web is maintained by LET-413. *Developmental biology* **268**, 448–56.

Botts, M. R., Cohen, L. B., Probert, C. S., Wu, F. and Troemel, E. R. (2016). Microsporidia Intracellular Development Relies on Myc Interaction Network Transcription Factors in the Host. *G3: Genes|Genomes|Genetics* **6**, 2707–2716.

Bowerman, B., Eaton, B. A. and Priess, J. R. (1992). skn-1, a maternally expressed gene required to specify the fate of ventral blastomeres in the early C. elegans embryo. *Cell* **68**, 1061–1075.

Bowman, S. K., Rolland, V., Betschinger, J., Kinsey, K. A., Emery, G. and Knoblich, J. A. (2008). The Tumor Suppressors Brat and Numb Regulate Transit-Amplifying Neuroblast Lineages in Drosophila. *Dev Cell* **14**, 535–546.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. *Genetics* **77**, 71–94.

Brock, T. J., Browse, J. and Watts, J. L. (2006). Genetic Regulation of Unsaturated Fatty

Acid Composition in C. elegans. *Plos Genet* **2**, e108.

Broday, L. (2017). The SUMO system in Caenorhabditis elegans development. *The International Journal of Developmental Biology* **61**, 159–164.

Broday, L., Kolotuev, I., Didier, C., Bhoumik, A., Gupta, B. P., Sternberg, P. W., Podbilewicz, B. and Ronai, Z. (2004). The small ubiquitin-like modifier (SUMO) is required for gonadal and uterine-vulval morphogenesis in Caenorhabditis elegans. *Genes & Development* **18**, 2380–2391.

Byrd, D. T. and Kimble, J. (2009). Scratching the niche that controls Caenorhabditis elegans germline stem cells. *Semin Cell Dev Biol* **20**, 1107–1113.

Cahan, P., Li, H., Morris, S. A., Lummertz da Rocha, E., Daley, G. Q. and Collins, J. J. (2014). CellNet: Network Biology Applied to Stem Cell Engineering. *Cell* **158**, 903–915.

Carrozza, M. J., Utley, R. T., Workman, J. L. and Côté, J. (2003). The diverse functions of histone acetyltransferase complexes. *Trends in Genetics* **19**, 321–329.

Ceol, C. J. and Horvitz, H. (2004). A new class of C. elegans synMuv genes implicates a Tip60/NuA4-like HAT complex as a negative regulator of Ras signaling. *Developmental cell* **6**, 563–76.

Ceron, J., Rual, J.-F., Chandra, A., Dupuy, D., Vidal, M. and van den Heuvel, S. (2007). Large-scale RNAi screens identify novel genes that interact with the C. elegans retinoblastoma pathway as well as splicing-related components with synMuv B activity. *BMC Developmental Biology*.

Chalfie, M. (1995). GREEN FLUORESCENT PROTEIN. *Photochem Photobiol* **62**, 651–656.

Chambers, S. M. and Studer, L. (2011). Cell Fate Plug and Play: Direct Reprogramming and Induced Pluripotency. *Cell* **145**, 827–830.

Chanda, S., Ang, C., Davila, J., Pak, C., Mall, M., Lee, Q., Ahlenius, H., Jung, S., Südhof, T. C. and Wernig, M. (2014). Generation of Induced Neuronal Cells by the Single Reprogramming Factor ASCL1. *Stem Cell Rep* **3**, 282–296.

Chang, W., Tilmann, C., Thoemke, K., Markussen, F.-H., Mathies, L. D., Kimble, J. and Zarkower, D. (2004). A forkhead protein controls sexual identity of the C. elegans male somatic gonad. *Development* **131**, 1425–1436.

Cheloufi, S. and Hochedlinger, K. (2017). Emerging roles of the histone chaperone CAF-1 in cellular plasticity. *Current Opinion in Genetics & Development* **46**, 83–94.

Chen, D. and Riddle, D. L. (2008). Function of the PHA-4/FOXA transcription factor

during C. eleganspost-embryonic development. *BMC Developmental Biology* **8**, 1–10.

Cinar, H., Richards, K. L., Oommen, K. S. and Newman, A. P. (2003). The EGL-13 SOX domain transcription factor affects the uterine pi cell lineages in Caenorhabditis elegans. *Genetics* **165**, 1623–8.

Ciosk, R., DePalma, M. and Priess, J. R. (2006). Translational Regulators Maintain Totipotency in the Caenorhabditis elegans Germline. *Science* **311**, 851–3.

Cirillo, L. A., Lin, F. R., Cuesta, I., Friedman, D., Jarnik, M. and Zaret, K. S. (2002). Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Molecular cell* **9**, 279–89.

Cooke, J. P., Sayed, N., Lee, J. and Wong, W. (2014). Innate immunity and epigenetic plasticity in cellular reprogramming. *Current Opinion in Genetics & Development* **28**, 89–91.

Cui, M. and Han, M. (2007). Roles of chromatin factors in C. elegans development. *WormBook : the online review of C. elegans biology* 1–16.

Cusanovich, D. A., Hill, A. J., Aghamirzaie, D., Daza, R. M., Pliner, H. A., Berletch, J. B., Filippova, G. N., Huang, X., Christiansen, L., DeWitt, W. S., et al. (2018a). A Single-Cell Atlas of In Vivo Mammalian Chromatin Accessibility. *Cell* **174**, 1309-1324.e18.

Cusanovich, D. A., Reddington, J. P., Garfield, D. A., Daza, R. M., Aghamirzaie, D., Marco-Ferreres, R., Pliner, H. A., Christiansen, L., Qiu, X., Steemers, F. J., et al. (2018b). The cis-regulatory dynamics of embryonic development at single-cell resolution. *Nature* **555**, 538–542.

Custer, L. M., Snyder, M. J., Flegel, K. and Csankovszki, G. (2014). The onset of C. elegans dosage compensation is linked to the loss of developmental plasticity. *Developmental biology* **385**, 279–90.

Dang-Nguyen, T. and Torres-Padilla, M.-E. (2015). How cells build totipotency and pluripotency: nuclear, chromatin and transcriptional architecture. *Curr Opin Cell Biol* **34**, 9–15.

Deppe, U., Schierenberg, E., Cole, T., Krieg, C., Schmitt, D., Yoder, B. and von Ehrenstein, G. (1978). Cell lineages of the embryo of the nematode Caenorhabditis elegans. *Proc National Acad Sci* **75**, 376–380.

Dineen, A., Nishimura, E., Goszczynski, B., Rothman, J. H. and McGhee, J. D. (2018). Quantitating transcription factor redundancy: The relative roles of the ELT-2 and ELT-7 GATA factors in the C. elegans endoderm. *Dev Biol* **435**, 150–161.

Djabrayan, N., Dudley, N. R., Sommermann, E. M. and Rothman, J. H. (2012). Essential role for Notch signaling in restricting developmental plasticity. *Genes & Development* **26**,

2386–2391.

Dodemont, H., Riemer, D., Ledger, N. and Weber, K. (1994). Eight genes and alternative RNA processing pathways generate an unexpectedly large diversity of cytoplasmic intermediate filament proteins in the nematode Caenorhabditis elegans. *Embo J* **13**, 2625–38.

Donaghey, J., Thakurela, S., Charlton, J., Chen, J. S., Smith, Z. D., Gu, H., Pop, R., Clement, K., Stamenova, E. K., Karnik, R., et al. (2018). Genetic determinants and epigenetic effects of pioneer-factor occupancy. *Nat Genet* **50**, 250–258.

Doppler, S. A., Deutsch, M.-A. A., Lange, R. and Krane, M. (2015). Direct Reprogramming-The Future of Cardiac Regeneration? *International journal of molecular sciences* **16**, 17368–93.

Doyon, Y., Selleck, W., Lane, W. S., Tan, S. and Côté, J. (2004). Structural and Functional Conservation of the NuA4 Histone Acetyltransferase Complex from Yeast to Humans. *Molecular and Cellular Biology* **24**, 1884–1896.

Drabikowski, K., Ferralli, J., Kistowski, M., Oledzki, J., Dadlez, M. and Chiquet-Ehrismann, R. (2018). Comprehensive list of SUMO targets in Caenorhabditis elegans and its implication for evolutionary conservation of SUMO signaling. *Scientific Reports* **8**, 1139.

Draper, B. W., Mello, C. C., Bowerman, B., Hardin, J. and Priess, J. R. (1996). MEX-3 Is a KH Domain Protein That Regulates Blastomere Identity in Early C. elegans Embryos. *Cell* **87**, 205–216.

Du, Z., He, F., Yu, Z., Bowerman, B. and Bao, Z. (2015). E3 ubiquitin ligases promote progression of differentiation during C. elegans embryogenesis. *Developmental biology* **398**, 267–79.

elegans Consortium, T. C. (2012). Large-Scale Screening for Targeted Knockouts in the Caenorhabditis elegans Genome. *G3 Genes Genomes Genetics* **2**, 1415–1425.

Ebrahimi, B. (2017). In vivo reprogramming for heart regeneration: A glance at efficiency, environmental impacts, challenges and future directions. *Journal of Molecular and Cellular Cardiology* **108**, 61–72.

Erdelyi, P., Wang, X., Suleski, M. and Wicky, C. (2017). A Network of Chromatin Factors Is Regulating the Transition to Postembryonic Development in Caenorhabditis elegans. *G3 (Bethesda, Md.)* **7**, 343–353.

Esteban, M. A. and Wang, J. (2017). Editorial overview: Cell reprogramming: Carpe diem. *Curr Opin Genet Dev* **46**, iv–vi.

Evans, T. C., Crittenden, S. L., Kodoyianni, V. and Kimble, J. (1994). Translational control of maternal glp-1 mRNA establishes an asymmetry in the C. elegans embryo. *Cell*

77, 183–194.

Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., Haw, R., Jassal, B., Korninger, F., May, B., et al. (2017). The Reactome Pathway Knowledgebase. *Nucleic Acids Res* **46**, gkx1132-.

Fakhouri, T. H., Stevenson, J., Chisholm, A. D. and Mango, S. E. (2010). Dynamic chromatin organization during foregut development mediated by the organ selector gene PHA-4/FoxA. *PLoS genetics* **6**,.

Fang, Z. and Cui, X. (2011). Design and validation issues in RNA-seq experiments. *Brief Bioinform* **12**, 280–287.

Firas, J. and Polo, J. M. (2017). Towards understanding transcriptional networks in cellular reprogramming. *Current Opinion in Genetics & Development* **46**, 1–8.

Fukushige, T. and Krause, M. (2005). The myogenic potency of HLH-1 reveals wide-spread developmental plasticity in early C. elegans embryos. *Development* **132**, 1795–805.

Fukushige, T., Hawkins, M. G. and McGhee, J. D. (1998). The GATA-factor elt-2 is essential for formation of the Caenorhabditis elegans intestine. *Developmental Biology* **198**, 286–302.

Gaiser, A. M., Brandt, F. and Richter, K. (2009). The Non-canonical Hop Protein from Caenorhabditis elegans Exerts Essential Functions and Forms Binary Complexes with Either Hsc70 or Hsp90. *J Mol Biol* **391**, 621–634.

Gao, F., Li, J., Zhang, H., Yang, X. and An, T. (2017). Identifying Candidate Reprogramming Genes in Mouse Induced Pluripotent Stem Cells. *Stem Cell Reviews and Reports* 1–10.

Gaspar-Maia, A., Alajem, A., Meshorer, E. and Ramalho-Santos, M. (2011). Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Bio* **12**, 36–47.

Gaudet, J. and Mango, S. (2002). Regulation of organogenesis by the Caenorhabditis elegans FoxA protein PHA-4. *Science (New York, N.Y.)* **295**, 821–5.

Gdula, D. A., Sandaltzopoulos, R., Tsukiyama, T., Ossipow, V. and Wu, C. (1998). Inorganic pyrophosphatase is a component of the Drosophila nucleosome remodeling factor complex. *Gene Dev* **12**, 3206–3216.

Ghosh, S. and Sternberg, P. W. (2014). Spatial and molecular cues for cell outgrowth during C. elegans uterine development. *Developmental biology*.

Giammartino, D. and Apostolou, E. (2016). The Chromatin Signature of Pluripotency: Establishment and Maintenance. *Curr Stem Cell Reports* **2**, 255–262.

Gilleard, J. and McGhee, J. (2001). Activation of hypodermal differentiation in the Caenorhabditis elegans embryo by GATA transcription factors ELT-1 and ELT-3. *Molecular and cellular biology* **21**, 2533–44.

Giroux, V. and Rustgi, A. K. (2017). Metaplasia: tissue injury adaptation and a precursor to the dysplasia–cancer sequence. *Nat Rev Cancer* **17**, 594–604.

Götz, M. and Jarriault, S. (2017). Programming and reprogramming the brain: a meeting of minds in neural fate. *Development* **144**, 2714–2718.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W. and Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biol* **59**, 307–321.

Guo, C. and Morris, S. A. (2017). Engineering cell identity: establishing new gene regulatory and chromatin landscapes. *Current Opinion in Genetics & Development* **46**, 50–57.

Hajduskova, M., Ahier, A., Daniele, T. and Jarriault, S. (2012). Cell plasticity in Caenorhabditis elegans: from induced to natural cell reprogramming. *Genesis (New York, N.Y. : 2000)* **50**, 1–17.

Hajduskova, M., Baytek, G., Kolundzic, E., Gosdschan, A., Kazmierczak, M., Ofenbauer, A., del Rosal, M., Herzog, S., ul Fatima, N., Mertins, P., et al. (2018). MRG-1/MRG15 Is a Barrier for Germ Cell to Neuron Reprogramming in Caenorhabditis elegans. *Genetics* **211**, genetics.301674.2018.

Halder, G., Callaerts, P. and Gehring, W. (1995). Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila. *Science* **267**, 1788–1792.

Hashimshony, T., Feder, M., Levin, M., Hall, B. K. and Yanai, I. (2015). Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. *Nature* **519**, 219–22.

Hatzold, J. and Conradt, B. (2008). Control of Apoptosis by Asymmetric Cell Division. *Plos Biol* **6**, e84.

Hemberger, M., Dean, W. and Reik, W. (2009). Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nat Rev Mol Cell Bio* **10**, 526–537.

Heng, J.-C. D., Orlov, Y. L. and Ng, H.-H. (2010). Transcription Factors for the Modulation of Pluripotency and Reprogramming. *Cold Spring Harb Sym* **75**, 237–244.

Herman, M. A. (2006). Hermaphrodite cell-fate specification. *WormBook : the online review of C. elegans biology* 1–16.

Hiatt, S. M., ren, H., Shyu, J. Y., Ellis, R. E., Hisamoto, N., Matsumoto, K., Kariya, K.-I., Kerppola, T. K. and Hu, C.-D. (2009). Caenorhabditis elegans FOS-1 and JUN-1 regulate plc-1 expression in the spermatheca to control ovulation. *Molecular biology of the cell* **20**, 3888–95.

Hope, I. A., Mounsey, A., Bauer, P. and Aslam, S. (2003). The forkhead gene family of Caenorhabditis elegans. *Gene* **304**, 43–55.

Horner, M. A., Quintin, S., Domeier, M., Kimble, J., Labouesse, M. and Mango, S. E. (1998). pha-4, anHNF-3 homolog, specifies pharyngeal organ identity inCaenorhabditis elegans. *Genes & Development* **12**, 1947–52.

Huber, P., Crum, T., Clary, L. M., Ronan, T., Packard, A. V. and Okkema, P. G. (2013). Function of the C. elegans T-box factor TBX-2 depends on SUMOylation. *Cell Mol Life Sci* **70**, 4157–4168.

Hüsken, K., Wiesenfahrt, T., Abraham, C., Windoffer, R., Bossinger, O. and Leube, R. E. (2008). Maintenance of the intestinal tube in Caenorhabditis elegans: the role of the intermediate filament protein IFC-2. *Differentiation* **76**, 881–896.

Hutter, H. and Schnabel, R. (1994). glp-1 and inductions establishing embryonic axes in C. elegans. *Dev Camb Engl* **120**, 2051–64.

Ichida, J. K., TCW, J., Julia, T., Williams, L. A., Carter, A. C., Shi, Y., Moura, M. T., Ziller, M., Singh, S., Amabile, G., et al. (2014). Notch inhibition allows oncogene-independent generation of iPS cells. *Nature Chemical Biology* **10**, 632–9.

Jahnel, O., Hoffmann, B., Merkel, R., Bossinger, O. and Leube, R. E. (2016). Chapter Twenty-Five Mechanical Probing of the Intermediate Filament-Rich Caenorhabditis Elegans Intestine. *Methods in Enzymology* **568**, 681–706.

Jarriault, S., Schwab, Y. and Greenwald, I. (2008). A Caenorhabditis elegans model for epithelial–neuronal transdifferentiation. *Proceedings of the National Academy of Sciences* **105**, 3790–5.

Johnson, E. S. (2004). PROTEIN MODIFICATION BY SUMO. *Annual Review of Biochemistry* **73**, 355–382.

Joshi, P. M., Riddle, M. R., Djabrayan, N. and Rothman, J. H. (2010). Caenorhabditis elegans as a model for stem cell biology. *Developmental Dynamics* **239**, 1539–54.

Kadyk, L. and Kimble, J. (1998). Genetic regulation of entry into meiosis in Caenorhabditis elegans. *Dev Camb Engl* **125**, 1803–13.

Kagias, K., Ahier, A., Fischer, N. and Jarriault, S. (2012). Members of the NODE (Nanog

and Oct4-associated deacetylase) complex and SOX-2 promote the initiation of a natural cellular reprogramming event in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 6596–601.

Kalb, J., Lau, K., Goszczynski, B., Fukushige, T., Moons, D., Okkema, P. and McGhee, J. (1998). pha-4 is Ce-fkh-1, a fork head/HNF-3alpha,beta,gamma homolog that functions in organogenesis of the C. elegans pharynx. *Development (Cambridge, England)* **125**, 2171–80.

Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Bot, N., Moreno, S., Sohrmann, M., et al. (2003). Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. *Nature* **421**, 231–237.

Kaminsky, R., Denison, C., Bening-Abu-Shach, U., Chisholm, A. D., Gygi, S. P. and Broday, L. (2009). SUMO Regulates the Assembly and Function of a Cytoplasmic Intermediate Filament Protein in C. elegans. *Developmental Cell* **17**, 724–735.

Karabinos, A., Schmidt, H., Harborth, J., Schnabel, R. and Weber, K. (2001). Essential roles for four cytoplasmic intermediate filament proteins in Caenorhabditis elegans development. *Proc National Acad Sci* **98**, 7863–7868.

Karabinos, A., Schulze, E., Schünemann, J., Parry, D. and Weber, K. (2003). In Vivo and in Vitro Evidence that the Four Essential Intermediate Filament (IF) Proteins A1, A2, A3 and B1 of the Nematode Caenorhabditis elegans Form an Obligate Heteropolymeric IF System. *J Mol Biol* **333**, 307–319.

Keeling, P. J. and Fast, N. M. (2002). MICROSPORIDIA: Biology and Evolution of Highly Reduced Intracellular Parasites. *Microbiology+* **56**, 93–116.

Kim, W., Underwood, R. S., Greenwald, I. and Shaye, D. D. (2018). OrthoList 2: A New Comparative Genomic Analysis of Human and Caenorhabditiselegans Genes. *Genetics* **210**, 445–461.

Kimble, J. and Crittenden, S. L. (2007). Controls of Germline Stem Cells, Entry into Meiosis, and the Sperm/Oocyte Decision in Caenorhabditis elegans. *Annu Rev Cell Dev Bi* **23**, 405–433.

Kimble, J. E. and White, J. G. (1981). On the control of germ cell development in Caenorhabditis elegans. *Dev Biol* **81**, 208–219.

Ko, K., Lee, W., Yu, J.-R. and Ahnn, J. (2007). PYP-1, inorganic pyrophosphatase, is required for larval development and intestinal function in C. elegans. *FEBS Letters* **581**, 5445–5453.

Koh, F., Sachs, M., Guzman-Ayala, M. and Ramalho-Santos, M. (2010). Parallel gateways to pluripotency: open chromatin in stem cells and development. *Curr Opin Genet Dev* **20**, 492–499.

Kolundzic, E., Ofenbauer, A., Bulut, S. I., Uyar, B., Baytek, G., Sommermeier, A., Seelk, S., He, M., Hirsekorn, A., Vucicevic, D., et al. (2018). FACT Sets a Barrier for Cell Fate Reprogramming in Caenorhabditis elegans and Human Cells. *Dev Cell* **46**, 611-626.e12.

Kudron, M. M., Victorsen, A., Gevirtzman, L., Hillier, L. W., Fisher, W. W., Vafeados, D., Kirkey, M., Hammonds, A. S., Gersch, J., Ammouri, H., et al. (2017). The modERN Resource: Genome-Wide Binding Profiles for Hundreds of Drosophila and Caenorhabditis elegans Transcription Factors. *Genetics* **208**, genetics.300657.2017.

Kurotsu, S., Suzuki, T. and Ieda, M. (2017). Direct Reprogramming, Epigenetics, and Cardiac Regeneration. *Journal of Cardiac Failure* **23**, 552–557.

Kwon, A. T., Arenillas, D. J., Hunt, R. and Wasserman, W. W. (2012). oPOSSUM-3: Advanced Analysis of Regulatory Motif Over-Representation Across Genes or ChIP-Seq Datasets. *G3 Genes Genomes Genetics* **2**, 987–1002.

Langley, A. R., Smith, J. C., Stemple, D. L. and Harvey, S. A. (2014). New insights into the maternal to zygotic transition. *Development* **141**, 3834–3841.

Lee, J., Sayed, N., Hunter, A., Au, K., Wong, W. H., Mocarski, E. S., Pera, R., Yakubov, E. and Cooke, J. P. (2012). Activation of Innate Immunity Is Required for Efficient Nuclear Reprogramming. *Cell* **151**, 547–558.

Lee, M.-H. and Schedl, T. (2006). RNA-binding proteins. *Wormbook Online Rev C Elegans Biology* 1–13.

Lei, R., Ye, K., Gu, Z. and Sun, X. (2015). Diminishing returns in next-generation sequencing (NGS) transcriptome data. *Gene* **557**, 82–7.

Levin, M., Hashimshony, T., Wagner, F. and Yanai, I. (2012). Developmental Milestones Punctuate Gene Expression in the Caenorhabditis Embryo. *Dev Cell* **22**, 1101–1108.

Levin, M., Anavy, L., Cole, A. G., Winter, E., Mostov, N., Khair, S., Senderovich, N., Kovalev, E., Silver, D. H., Feder, M., et al. (2016). The mid-developmental transition and the evolution of animal body plans. *Nature* **531**, 637–641.

Liu, Y., Zhou, J. and White, K. P. (2014). RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics (Oxford, England)* **30**, 301–4.

Lomelí, H. and Vázquez, M. (2011). Emerging roles of the SUMO pathway in development. *Cellular and Molecular Life Sciences* **68**, 4045–4064.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550.

Luallen, R. J., Bakowski, M. A. and Troemel, E. R. (2015). Characterization of microsporidia-induced developmental arrest and a transmembrane leucine-rich repeat protein in Caenorhabditis elegans. *PloS one* **10**, e0124065.

MacNeil, L. T., Pons, C., Arda, H., Giese, G. E., Myers, C. L. and Walhout, A. J. (2015). Transcription Factor Activity Mapping of a Tissue-Specific in vivo Gene Regulatory Network. *Cell systems* **1**, 152–162.

Maduro, M. F. (2009). Structure and evolution of the C. elegans embryonic endomesoderm network. *Biochimica et biophysica acta* **1789**, 250–60.

Maduro, M. F. (2015). Developmental robustness in the Caenorhabditis elegans embryo. *Molecular reproduction and development*.

Maduro, M. F., Hill, R. J., Heid, P. J., Newman-Smith, E. D., Zhu, J., Priess, J. R. and Rothman, J. H. (2005). Genetic redundancy in endoderm specification within the genus Caenorhabditis. *Developmental Biology* **284**, 509–22.

Maduro, M. F., Broitman-Maduro, G., Mengarelli, I. and Rothman, J. H. (2007). Maternal deployment of the embryonic SKN-1→MED-1,2 cell specification pathway in C. elegans. *Developmental Biology* **301**, 590–601.

Maduro, M. F., Broitman-Maduro, G., Choi, H., Carranza, F., Wu, A. C. and Rifkin, S. A. (2015). MED GATA factors promote robust development of the C. elegans endoderm. *Developmental biology* **404**, 66–79.

Maduro, M. F. and Rothman, J. H. (2002). Making worm guts: the gene regulatory network of the Caenorhabditis elegans endoderm. *Developmental biology* **246**, 68–85.

Mango, S., Lambie, E. and Kimble, J. (1994). The pha-4 gene is required to generate the pharyngeal primordium of Caenorhabditis elegans. *Development (Cambridge, England)* **120**, 3019–31.

Masaki, T., McGlinchey, A., Cholewa-Waclaw, J., Qu, J., Tomlinson, S. R. and Rambukkana, A. (2014). Innate Immune Response Precedes Mycobacterium leprae–Induced Reprogramming of Adult Schwann Cells. *Cell Reprogramming Former Cloning Stem Cells* **16**, 9–17.

McGhee, J. D., Fukushige, T., Krause, M. W., Minnema, S. E., Goszczynski, B., Gaudet, J., Kohara, Y., Bossinger, O., Zhao, Y., Khattra, J., et al. (2009). ELT-2 is the predominant transcription factor controlling differentiation and function of the C. elegans intestine, from embryo to adult. *Developmental Biology* **327**, 551–565.

Mello, C. C., Draper, B. W. and Prless, J. R. (1994). The maternal genes apx-1 and glp-1 and establishment of dorsal-ventral polarity in the early C. elegans embryo. *Cell* **77**, 95–106.

Merrell, A. J. and Stanger, B. Z. (2016). Adult cell plasticity in vivo: de-differentiation and transdifferentiation are back in style. *Nature Reviews Molecular Cell Biology* **17**, 413–425.

Merritt, C., Rasoloson, D., Ko, D. and Seydoux, G. (2008). 3′ UTRs Are the Primary Regulators of Gene Expression in the C. elegans Germline. *Curr Biol* **18**, 1476–1482.

Miyabayashi, T., Palfreyman, M. T., Sluder, A. E., Slack, F. and Sengupta, P. (1999). Expression and Function of Members of a Divergent Nuclear Receptor Family in Caenorhabditis elegans. *Dev Biol* **215**, 314–331.

Mootz, D., Ho, D. M. and Hunter, C. P. (2004). The STAR/Maxi-KH domain protein GLD-1 mediates a developmental switch in the translational control of C. elegans PAL-1. *Development* **131**, 3263–3272.

Morris, S. A., Cahan, P., Li, H., Zhao, A. M., San Roman, A. K., Shivdasani, R. A., Collins, J. J. and Daley, G. Q. (2014). Dissecting Engineered Cell Types and Enhancing Cell Fate Conversion via CellNet. *Cell* **158**, 889–902.

Moskowitz, I. and Rothman, J. (1996). lin-12 and glp-1 are required zygotically for early embryonic cellular interactions and are regulated by maternal GLP-1 signaling in Caenorhabditis elegans. *Dev Camb Engl* **122**, 4105–17.

Moskowitz, I., Gendreau, S. and Rothman, J. (1994). Combinatorial specification of blastomere identity by glp-1-dependent cellular interactions in the nematode Caenorhabditis elegans. *Development (Cambridge, England)* **120**, 3325–38.

Nance, J., Lee, J.-Y. and Goldstein, B. (2005). Gastrulation in C. elegans. *Wormbook* 1–13.

Neff, A. W., King, M. W. and Mescher, A. L. (2011). Dedifferentiation and the role of sall4 in reprogramming and patterning during amphibian limb regeneration. *Dev Dynam* **240**, 979–989.

Neumüller, R. A., Betschinger, J., Fischer, A., Bushati, N., Poernbacher, I., Mechtler, K., Cohen, S. M. and Knoblich, J. A. (2008). Mei-P26 regulates microRNAs and cell growth in the Drosophila ovarian stem cell lineage. *Nature* **454**, 241–245.

Newport, J. and Kirschner, M. (1982). A major developmental transition in early xenopus embryos: II. control of the onset of transcription. *Cell* **30**, 687–696.

O'Brien, D., Jones, L. M., Good, S., Miles, J., Vijayabaskar, Aston, R., Smith, C. E., Westhead, D. R. and van Oosten-Hawle, P. (2018). A PQM-1-Mediated Response Triggers Transcellular Chaperone Signaling and Regulates Organismal Proteostasis. *Cell Reports* **23**, 3905–3919.

Ofenbauer, A. and Tursun, B. (2019). Strategies for in vivo reprogramming. *Curr Opin Cell Biol* **61**, 9–15.

Okkema, P. and Fire, A. (1994). The Caenorhabditis elegans NK-2 class homeoprotein CEH-22 is involved in combinatorial activation of gene expression in pharyngeal muscle. *Dev Camb Engl* **120**, 2175–86.

Okkema, P., Harrison, S., Plunger, V., Aryana, A. and Fire, A. (1993). Sequence requirements for myosin gene expression and regulation in Caenorhabditis elegans. *Genetics* **135**, 385–404.

Okkema, P., Ha, E., Haun, C., Chen, W. and Fire, A. (1997). The Caenorhabditis elegans NK-2 homeobox gene ceh-22 activates pharyngeal muscle gene expression in combination with pha-1 and is required for normal pharyngeal development. *Dev Camb Engl* **124**, 3965–73.

Onder, T. T., Kara, N., Cherry, A., Sinha, A. U., Zhu, N., Bernt, K. M., Cahan, P., Mancarci, O. B., Unternaehrer, J., Gupta, P. B., et al. (2012). Chromatin-modifying enzymes as modulators of reprogramming. *Nature* **483**, 598–602.

Oommen, K. S. and Newman, A. P. (2007). Co-regulation by Notch and Fos is required for cell fate specification of intermediate precursors during C. elegans uterine development. *Dev Camb Engl* **134**, 3999–4009.

Pagano, J. M., Farley, B. M., Essien, K. I. and Ryder, S. P. (2009). RNA recognition by the embryonic cell fate determinant and germline totipotency factor MEX-3. *Proc National Acad Sci* **106**, 20252–20257.

Patel, T., Tursun, B., Rahe, D. P. and Hobert, O. (2012). Removal of Polycomb Repressive Complex 2 Makes C. elegans Germ Cells Susceptible to Direct Conversion into Specific Somatic Cell Types. *Cell Reports* **2**, 1178–86.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods* **14**, 417–419.

Peixoto, L., Risso, D., Poplawski, S. G., Wimmer, M. E., Speed, T. P., Wood, M. A. and Abel, T. (2015). How data analysis affects power, reproducibility and biological insight of RNA-seq studies in complex datasets. *Nucleic acids research* **43**, 7664–74.

Pereira, B., Borgne, M., Chartier, N. T., Billaud, M. and Almeida, R. (2013). MEX-3 proteins: recent insights on novel post-transcriptional regulators. *Trends Biochem Sci* **38**, 477–479.

Pohl, C. and Dikic, I. (2019). Cellular quality control by the ubiquitin-proteasome system and autophagy. *Science* **366**, 818–822.

Poillet-Perez, L. and White, E. (2019). Role of tumor and host autophagy in cancer

metabolism. *Gene Dev* **33**, 610–619.

Priess, J. R. and Thomson, J. N. (1987). Cellular interactions in early C. elegans embryos. *Cell* **48**, 241–250.

Priess, J. R., Schnabel, H. and Schnabel, R. (1987). The glp-1 locus and cellular interactions in early C. elegans embryos. *Cell* **51**, 601–611.

Qian, L., Huang, Y., Spencer, I. C., Foley, A., Vedantham, V., Liu, L., Conway, S. J., Fu, J. and Srivastava, D. (2012). In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. *Nature* **485**, 593.

Qin, H., Diaz, A., Blouin, L., Lebbink, R., Patena, W., Tanbun, P., LeProust, E. M., McManus, M. T., Song, J. S. and Ramalho-Santos, M. (2014). Systematic Identification of Barriers to Human iPSC Generation. *Cell* **158**, 449–461.

Quintin, S., Michaux, G., McMahon, L., Gansmuller, A. and Labouesse, M. (2001). The Caenorhabditis elegans Gene lin-26 Can Trigger Epithelial Differentiation without Conferring Tissue Specificity. *Developmental Biology* 410–421.

Raj, A., Rifkin, S. A., Andersen, E. and van Oudenaarden, A. (2010). Variability in gene expression underlies incomplete penetrance. *Nature* **463**, 913–8.

Rajan, M., Anderson, C. P., Rindler, P. M., Romney, S., dos Santos, M. C., Gertz, J. and Leibold, E. A. (2019). NHR-14 loss of function couples intestinal iron uptake with innate immunity in C. elegans through PQM-1 signaling. *Elife* **8**, e44674.

Ramani, V., Deng, X., Qiu, R., Gunderson, K. L., Steemers, F. J., steche, C., Noble, W. S., Duan, Z. and Shendure, J. (2017). Massively multiplex single-cell Hi-C. *Nat Methods* **14**, 263–266.

Rangaraju, S., lis, G., Thompson, R. C., Gomez-Amaro, R. L., Kurian, L., Encalada, S. E., Niculescu, A. B., Salomon, D. R. and Petrascheck, M. (2015). Suppression of transcriptional drift extends C. elegans lifespan by postponing the onset of mortality. *eLife* **4**,.

Reddy, K. C., Dror, T., Sowa, J. N., Panek, J., Chen, K., Lim, E. S., Wang, D. and Troemel, E. R. (2017). An Intracellular Pathogen Response Pathway Promotes Proteostasis in C. elegans. *Current Biology* **27**, 3544-3553.e5.

Reece-Hoyes, J. S., Shingles, J., Dupuy, D., Grove, C. A., Walhout, A. J., Vidal, M. and Hope, I. A. (2007). Insight into transcription factor gene duplication from Caenorhabditis elegans Promoterome-driven expression patterns. *BMC genomics* **8**, 27.

Rezvani, M., Español-Suñer, R., Malato, Y., Dumont, L., Grimm, A. A., Kienle, E., Bindman, J. G., Wiedtke, E., Hsu, B. Y., Naqvi, S. J., et al. (2016). In Vivo Hepatic Reprogramming of Myofibroblasts with AAV Vectors as a Therapeutic Strategy for Liver

Fibrosis. *Cell Stem Cell* **18**, 809–816.

Ricci, M., Cosma, M. and Lakadamyali, M. (2017). Super resolution imaging of chromatin in pluripotency, differentiation, and reprogramming. *Curr Opin Genet Dev* **46**, 186–193.

Richard, J., Zuryn, S., Fischer, N., Pavet, V., Vaucamps, N. and Jarriault, S. (2011). Direct in vivo cellular reprogramming involves transition through discrete, non-pluripotent steps. *Development* **138**, 1483–1492.

Riddle, M. R., Weintraub, A., Nguyen, K. C., Hall, D. H. and Rothman, J. H. (2013). Transdifferentiation and remodeling of post-embryonic C. elegans cells by a single transcription factor. *Development* **140**, 4844–4849.

Riddle, M. R., Spickard, E. A., Jevince, A., Nguyen, K., Hall, D. H., Joshi, P. M. and Rothman, J. H. (2016). Transorganogenesis and transdifferentiation in C. elegans are dependent on differentiated cell identity. *Developmental Biology* **420**, 136–147.

Robert, V. J., Mercier, M. G., Bedet, C., Janczarski, S., Merlet, J., Garvis, S., Ciosk, R. and Palladino, F. (2014). The SET-2/SET1 Histone H3K4 Methyltransferase Maintains Pluripotency in the Caenorhabditis elegans Germline. *Cell Reports* **9**, 443–450.

Robertson, S. and Lin, R. (2015). Current Topics in Developmental Biology. *Current topics in developmental biology* 1–42.

Rothman, J. and Jarriault, S. (2019). Developmental Plasticity and Cellular Reprogramming in Caenorhabditis elegans. *Genetics* **213**, 723–757.

Rual, J.-F., Ceron, J., Koreth, J., Hao, T., Nicot, A.-S., Hirozane-Kishikawa, T., Vandenhaute, J., Orkin, S. H., Hill, D. E., van den Heuvel, S., et al. (2004). Toward Improving Caenorhabditis elegans Phenome Mapping With an ORFeome-Based RNAi Library. *Genome Res* **14**, 2162–2168.

Ruan, J., Li, H., Chen, Z., Coghlan, A., Coin, L. M., Guo, Y., Hériché, J.-K., Hu, Y., Kristiansen, K., Li, R., et al. (2008). TreeFam: 2008 Update. *Nucleic Acids Res* **36**, D735–D740.

Sahu, S. N., Lewis, J., Patel, I., Bozdag, S., Lee, J. H., Sprando, R. and Cinar, H. (2013). Genomic Analysis of Stress Response against Arsenic in Caenorhabditis elegans. *Plos One* **8**, e66431.

Sammut, M., Cook, S. J., Nguyen, K. C., Felton, T., Hall, D. H., Emmons, S. W., Poole, R. J. and Barrios, A. (2015). Glia-derived neurons are required for sex-specific learning in C. elegans. *Nature* **526**, 385–390.

Sayed, N., Wong, W., Ospino, F., Meng, S., Lee, J., Jha, A., Dexheimer, P., Aronow, B. J. and Cooke, J. P. (2015). Transdifferentiation of Human Fibroblasts to Endothelial Cells.

Circulation **131**, 300–309.

Schier, A. F. (2007). The Maternal-Zygotic Transition: Death and Birth of RNAs. *Science* **316**, 406–407.

Schwamborn, J. C., Berezikov, E. and Knoblich, J. A. (2009). The TRIM-NHL Protein TRIM32 Activates MicroRNAs and Prevents Self-Renewal in Mouse Neural Progenitors. *Cell* **136**, 913–925.

Seelk, S., Adrian-Kalchhauser, I., Hargitai, B., Hajduskova, M., Gutnik, S., Tursun, B. and Ciosk, R. (2016). Increasing Notch signaling antagonizes PRC2-mediated silencing to promote reprograming of germ cells into neurons. *eLife* **5**,.

Seydoux, G., Salvage, C. and Greenwald, I. (1993). Isolation and Characterization of Mutations Causing Abnormal Eversion of the Vulva in Caenorhabditis elegans. *Dev Biol* **157**, 423–436.

Shapira, M., Hamlin, B. J., Rong, J., Chen, K., Ronen, M. and Tan, M.-W. W. (2006). A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 14086–91.

Shaye, D. D. and Greenwald, I. (2011). OrthoList: A Compendium of C. elegans Genes with Human Orthologs. *PLoS ONE* **6**, e20085.

Shpigel, N., Shemesh, N., Kishner, M. and Ben-Zvi, A. (2019). Dietary restriction and gonadal signaling differentially regulate post-development quality control functions in Caenorhabditis elegans. *Aging Cell* **18**, e12891.

Shu, M., Palas, C., Thandavarayan, R. A. and Cooke, J. P. (2018). Transflammation: How Innate Immune Activation and Free Radicals Drive Nuclear Reprogramming. *Antioxidants & Redox Signaling* **29**, 205–218.

Slack, J. (2007). Metaplasia and transdifferentiation: from pure biology to the clinic. *Nature reviews. Molecular cell biology* **8**, 369–78.

Slack, F. J., Basson, M., Liu, Z., Ambros, V., Horvitz, H. R. and Ruvkun, G. (2000). The lin-41 RBCC Gene Acts in the C. elegans Heterochronic Pathway between the let-7 Regulatory RNA and the LIN-29 Transcription Factor. *Mol Cell* **5**, 659–669.

Slaidina, M. and Lehmann, R. (2014). Translational control in germline stem cell development. *J Cell Biology* **207**, 13–21.

Smith, D. K., He, M., Zhang, C.-L. and Zheng, J. C. (2017). The therapeutic potential of cell identity reprogramming for the treatment of aging-related neurodegenerative disorders. *Progress in Neurobiology* **157**, 212–229.

Sommermann, E., Strohmaier, K. R., Maduro, M. F. and Rothman, J. H. (2010). Endoderm development in Caenorhabditis elegans: the synergistic action of ELT-2 and -7 mediates the specification→differentiation transition. *Developmental biology* **347**, 154–66.

Soufi, A., Donahue, G. and Zaret, K. S. (2012). Facilitators and Impediments of the Pluripotency Reprogramming Factors' Initial Engagement with the Genome. *Cell* **151**, 994–1004.

Spencer, W., Zeller, G., Watson, J. D., Henz, S. R., Watkins, K. L., McWhirter, R. D., Petersen, S., edharan, V. T., Widmer, C., Jo, J., et al. (2011). A spatial and temporal map of C. elegans gene expression. *Genome research* **21**, 325–41.

Spickard, E. A., Joshi, P. M. and Rothman, J. H. (2018). The multipotency-to-commitment transition in Caenorhabditis elegans—implications for reprogramming from cells to organs. *FEBS Letters*.

Srivastava, D. and DeWitt, N. (2016). In Vivo Cellular Reprogramming: The Next Generation. *Cell* **166**, 1386–1396.

Stainier, D. Y. (2005). No Organ Left Behind: Tales of Gut Development and Evolution. *Science* **307**, 1902–1904.

Stiernagle, T. (2006). Maintenance of C. elegans. *Wormbook Online Rev C Elegans Biology* 1–11.

Sui, S. J., Mortimer, J. R., Arenillas, D. J., Brumm, J., Walsh, C. J., Kennedy, B. P. and Wasserman, W. W. (2005). oPOSSUM: identification of over-represented transcription factor binding sites in co-expressed genes. *Nucleic Acids Res* **33**, 3154–3164.

Sui, S. J., Fulton, D. L., Arenillas, D. J., Kwon, A. T. and Wasserman, W. W. (2007). oPOSSUM: integrated tools for analysis of regulatory motif over-representation. *Nucleic Acids Res* **35**, W245–W252.

Sulston, J. E. and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Dev Biol* **56**, 110–156.

Sulston, J. E., Schierenberg, E., White, J. G. and Thomson, J. N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. *Developmental Biology* **100**, 64–119.

Tadros, W. and Lipshitz, H. D. (2009). The maternal-to-zygotic transition: a play in two acts. *Development* **136**, 3033–3042.

Taguchi, J. and Yamada, Y. (2017). In vivo reprogramming for tissue regeneration and organismal rejuvenation. *Current Opinion in Genetics & Development* **46**, 132–140.

Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–76.

Takahashi, K. and Yamanaka, S. (2016). A decade of transcription factor-mediated reprogramming to pluripotency. *Nature reviews. Molecular cell biology*.

Tawe, W. N., Eschbach, M.-L., Walter, R. D. and Henkle-Dührsen, K. (1998). Identification of stress-responsive genes in Caenorhabditis elegans using RT-PCR differential display. *Nucleic Acids Res* **26**, 1621–1627.

Thomas, N., Kenrick, M., Giesler, T., Kiser, G., Tinkler, H. and Stubbs, S. (2005). Characterization and gene expression profiling of a stable cell line expressing a cell cycle GFP sensor. *Cell Cycle Georget Tex* **4**, 191–5.

Tocchini, C., Keusch, J. J., Miller, S. B., Finger, S., Gut, H., Stadler, M. B. and Ciosk, R. (2014). The TRIM-NHL Protein LIN-41 Controls the Onset of Developmental Plasticity in Caenorhabditis elegans. *PLoS Genetics* **10**, e1004533.

Toh, C.-X. D., Chan, J.-W., Chong, Z.-S., Wang, H. F., Guo, H. C., Satapathy, S., Ma, D., Goh, G. Y. L., Khattar, E., Yang, L., et al. (2016). RNAi Reveals Phase-Specific Global Regulators of Human Somatic Cell Reprogramming. *Cell Reports* **15**, 2597–2607.

Torres-Padilla, M.-E. and Chambers, I. (2014). Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage. *Development* **141**, 2173–2181.

Troemel, E. R., Félix, M.-A., Whiteman, N. K., Barrière, A. and Ausubel, F. M. (2008). Microsporidia Are Natural Intracellular Parasites of the Nematode Caenorhabditis elegans. *PLoS Biology* 2736–52.

Tursun, B. (2012). Cellular reprogramming processes in Drosophila and C. elegans. *Current opinion in genetics & development* **22**, 475–84.

Tursun, B., Patel, T., Kratsios, P. and Hobert, O. (2011). Direct Conversion of C. elegans Germ Cells into Specific Neuron Types. *Science* **331**, 304–8.

van Attikum, H. and Gasser, S. M. (2005). The histone code at DNA breaks: a guide to repair? *Nature Reviews Molecular Cell Biology* **6**, 757–765.

Vierbuchen, T. and Wernig, M. (2012). Molecular Roadblocks for Cellular Reprogramming. *Molecular Cell* **47**, 827–838.

Wagner, A., Regev, A. and Yosef, N. (2016). Revealing the vectors of cellular identity with single-cell genomics. *Nat Biotechnol* **34**, 1145–1160.

Ward, J. D., Bojanala, N., Bernal, T., Ashrafi, K., Asahina, M. and Yamamoto, K. R. (2013). Sumoylated NHR-25/NR5A Regulates Cell Fate duringC. elegans Vulval

Development. *PLoS Genetics* **9**, e1003992.

Weintraub, H., Tapscott, S., Davis, R., Thayer, M., Adam, M., Lassar, A. and Miller, A. (1989). Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc National Acad Sci* **86**, 5434–5438.

Wiesenfahrt, T., Berg, J. Y., Nishimura, E., Robinson, A. G., Goszczynski, B., Lieb, J. D. and McGhee, J. D. (2016). The function and regulation of the GATA factor ELT-2 in the C. elegans endoderm. *Development (Cambridge, England)* **143**, 483–91.

Xu, J., Du, Y. and Deng, H. (2015). Direct Lineage Reprogramming: Strategies, Mechanisms, and Applications. *Cell Stem Cell* **16**, 119–134.

Yan, B., Memar, N., Gallinger, J. and Conradt, B. (2013). Coordination of Cell Proliferation and Cell Fate Determination by CES-1 Snail. *Plos Genet* **9**, e1003884.

Yang, W., Dierking, K. and Schulenburg, H. (2016). WormExp: a web-based application for a Caenorhabditis elegans -specific gene expression enrichment analysis. *Bioinformatics* **32**, 943–945.

Yu, G., Wang, L.-G., Han, Y. and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omics J Integr Biology* **16**, 284–287.

Yuzyuk, T., Fakhouri, T., Kiefer, J. and Mango, S. (2009). The polycomb complex protein mes-2/E(z) promotes the transition from developmental plasticity to differentiation in C. elegans embryos. *Developmental cell* **16**, 699–710.

Zahler, A. M. (2012). Pre-mRNA splicing and its regulation in Caenorhabditis elegans. *Wormbook Online Rev C Elegans Biology* 1–21.

Zhang, M., Lin, Y.-H., Sun, Y., Zhu, S., Zheng, J., Liu, K., Cao, N., Li, K., Huang, Y. and Ding, S. (2016). Pharmacological Reprogramming of Fibroblasts into Neural Stem Cells by Signaling-Directed Transcriptional Activation. *Cell Stem Cell* **18**, 653–667.

Zhang, H., Smolen, G. A., Palmer, R., Christoforou, A., van den Heuvel, S. and Haber, D. A. (2004). SUMO modification is required for in vivo Hox gene regulation by the Caenorhabditis elegans Polycomb group protein SOP-2. *Nature Genetics* 507–11.

Zhong, M., Niu, W., Lu, Z. J., Sarov, M., Murray, J. I., Janette, J., Raha, D., Sheaffer, K. L., Lam, H. Y., Preston, E., et al. (2010). Genome-wide identification of binding sites defines distinct functions for Caenorhabditis elegans PHA-4/FOXA in development and environmental response. *PLoS genetics* **6**, e1000848.

Zhu, J., Fukushige, T., McGhee, J. and Rothman, J. (1998). Reprogramming of early embryonic blastomeres into endodermal progenitors by a Caenorhabditis elegans GATA factor. *Genes & development* **12**, 3809–14.

Zhu, J., Hill, R., Heid, P., Fukuyama, M., Sugimoto, A., Priess, J. and Rothman, J. (1997). end-1 encodes an apparent GATA factor that specifies the endoderm precursor in Caenorhabditis elegans embryos. *Genes & development* **11**, 2883–96.

Zimmerman, S. M., Hinkson, I. V., Elias, J. E. and Kim, S. K. (2015). Reproductive Aging Drives Protein Accumulation in the Uterus and Limits Lifespan in C. elegans. *PLOS Genetics* **11**, e1005725.

Zuryn, S., Daniele, T. and Jarriault, S. (2012). Direct cellular reprogramming in Caenorhabditis elegans: facts, models, and promises for regenerative medicine. *Wiley Interdisciplinary Reviews: Developmental Biology* **1**, 138–152.

Zuryn, S., Ahier, A., Portoso, M., White, E. R., Morin, M.-C. C., Margueron, R. and Jarriault, S. (2014). Transdifferentiation. Sequential histone-modifying activities determine the robustness of transdifferentiation. *Science (New York, N.Y.)* **345**, 826–9.